LAND AND WATER INFORMATION STANDARDS IN NEW BRUNSWICK

A Report Submitted to the
Surveys and Mapping Committee
Province of New Brunswick

LEGISLATIVE LIBRARY
BIBLIOTHEQUE DE L'ASSEMBLÉE
LEGISLATIVE

FEB 07 1990

FREDERICTON, N.B.

Prepared by:

THE CABOT GROUP INC

in association with
THE UNIVERSITY OF NEW BRUNSWICK
and
UNIVERSAL SYSTEMS LTD.

March, 1989

90-02662/

LAND AND WATER INFORMATION STANDARDS IN NEW BRUNSWICK

EXECUTIVE SUMMARY

This report presents the findings of a project that investigated land and water information standards in New Brunswick. It is an extension of efforts to develop a land information network for the Province, as reported in the Final Report of the Landnet Program, Phase 1 (1988). Specific objectives of this standards project included:

- developing a standards framework for recording, classifying, coding and storing data, related to land and water management in New Brunswick.
- preparing an inventory of existing standards based on the above framework.
- identifying any conflicts, inconsistencies or major omissions in current standards which should be addressed by the Provincial Government.

BACKGROUND

An integrated land information system relates different types of land information to a common base composed of: (i) a coordinate framework, (ii) a series of base maps and (iii) a cadastral overlay. These components comprise a "foundation sub-system" that allows the linking between other, more specialized themes collected by line departments (e.g., forest cover, soils, etc.). Traditionally, government departments have collected information in accordance with their own internal standards with little regard for the needs of outside users and with little notice of the types of data collected by these other departments. These attitudes, while understandable, have led to increasing duplication of effort among different organizations during the map- and database maintenance stages. While this duplication and redundancy appears to have no effect on individual departmental budgets, the overall costs to the province are substantial — resulting in increasing interest in managing provincial land- and water-related information from a *corporate* rather than departmental perspective.

Much of the data collected and mapped by different departments is specialized in nature; e.g., mineral deposits, forest cover, interpolated soil samples, well locations and underground storage facilities. At the same time, most of these same departments also share a common need for information from base maps, property maps and property ownership records. In fact, some users update or correct such information (for their own purposes, at least) when these records are found to be out-of-date or erroneous. However, because of the different standards employed in this process, other users often cannot make direct use of this updated information and must revise these same records themselves as the need arises. Further, in collecting information from different departments, members of the general public must deal with information portrayed to several different standards of symbology, reliability and timeliness.

While it is too late to resolve differences among existing hardcopy maps and records in different provincial departments, it is not too late to develop common standards for digital information. Such standards would satisfy three objectives:

- TRANSPARENCY: They would allow different departments to access, interpret and compile data residing on different information systems in different departments. Ideally, this process would be *transparent* to the end-user — it would appear as though all the information was coming directly from the department's own system.
- CONSISTENCY: They would enable the province to accept data from contractors, share data among different departments and distribute information to the general public according to a single set of specifications.
- DEFAULT EXAMPLES: They would provide a collection of "default" specifications or examples which could be used by provincial organizations just acquiring the technology and beginning their own programs.

STANDARDS DEFINITION

For the purposes of this project a "standard" was defined as *a specification which describes a dataset (conventional or digital) commonly employed by one or more user groups outside the agency or department which produces that dataset.* Examples of standards which fit this definition include:

DOT specifications for road classification;

Reference ellipsoids for basic definition of latitude and longitude;

Map projections for spatial referencing of point information;

LRIS, EMR and DNR specifications for digital base mapping;

LRIS, PPDC and Assessment Branch parcel definitions and their corresponding identifier numbers;

WHY STANDARDS ?

- Supplemented by the appropriate political and institutional mechanisms, standards can provide the technical basis for limiting the amount of duplication in collection, identification, verification and maintenance of land and water information in New Brunswick.

- If implemented effectively standards are a means of protecting and even maximizing the Province's investment in digital data.

- Standards will facilitate greater distribution of databases to the general public and thereby increase potential revenue through the value-added reselling of information.

RESULTS OF SITE INTERVIEWS

Although this study identified very few official standards in New Brunswick based on the site interviews, those that were identified could be categorized as follows:

- CATEGORY 1: Some databases were developed with the specific intention of creating a consistent, Province-wide information base general use (eg., LRIS base mapping and Parcel Index). Although not officially referred to as a standard, these databases have detailed documentation available concerning content, lineage of information, and format.

- CATEGORY 2: Other major databases became popular simply because:
  - The information contained in them was found to be reasonably up-to-date and accurate;
  - The database coverage spanned most if not all of the Province; and,
  - The database could be readily accessed by outside users.

These databases became *de-facto* standards because databases that were originally designed as standards were not complete or were not easily accessible by an outside departments. Examples of such standards include the DNR 20-chain mapping, the NTS 1:50,000 map series and the PATS assessment database.

From the site interviews, it is also clear that most standards that do exist are in the "foundation sub-system", as described above. Common standards for data definition, classification, representation and communication of information are not prevalent among Provincial departments.

Two recurring themes that emerged from the site interview process involved:

- DATABASE DOCUMENTATION. With few exceptions, agencies which developed their information systems in-house and now use them exclusively for internal purposes, generally could not provide detailed documentation covering data collection, maintenance and reliability. Agencies involved in contract management and/or data transfer with other organizations generally possessed more detailed production specifications and database documentation.

- UPDATING OF DATABASES. In some cases, databases are being built initially using the best information available at the time of compilation. Because emphasis is usually placed on obtaining Province-wide (or project-wide) coverage as efficiently as possible, the resulting reliability of may be relatively low during early stages. In most textual databases examined, basic ownership, address and parcel information is collected through program application forms and is maintained on an exception-basis only. Since LRIS and EMR base mapping is updated so infrequently, there were examples of project-specific map revision in virtually every map-using department interviewed. Updating in this manner results in 'pockets' of the database being more accurate or up-to-date than others. This inconsistency has been recognized as a serious concern by experienced database users that have to deal with information from multiple sources.

Site interviews highlighted the desire of many users to have an easily-accessible, single point of contact for the dissemination of up-to-date parcel ownership, land cover and land use information. In a land information network, while database queries may be made through a single source, the necessary data may be retrieved from many different systems. In its envisioned form, the user would specify or request information related to digital maps, records or images based on his/her own "view of the world" in terms of feature definition, representation ,etc. Similarly, output from the network would be defined and represented in accordance with the user's own specifications and practices. Ideally —

whether the required information is gathered from the database at hand or retrieved from a number of different outside systems — *the whole process should be transparent to the end-user* - a process made possible through the use of standards.

PROPOSED STANDARDS

Adoption or development of the following standards is proposed:

- **Standard for Relevant Terminology.** A standard Provincial set of relevant terms and definitions for land- and water-related information management is necessary to facilitate better understanding of the issues and to reach consensus on the detailed planning of standards.
- **Standard for Spatial Referencing.** A complete spatial referencing standard provides the practical basis for explicitly locating, recording and retrieving features on the ground. More importantly, it provides the mathematical basis for consistent digital coverage across the Province for use within GIS applications.
- **Standard for Parcel Definition & Identification.** While Provincial departments will continue to use management parcels of different sizes and composition, there must be some means of relating these parcels to one another on a one-to-one, many-to-one, or one-to-many basis.
- **Standard for Feature or Entity Classification**. Integrating information — or simply determining the suitability of information available from various sources — requires a common framework for the classification of basic and thematic information.
- **Standard for Object Representation**. Standards for object representation govern how features are to be collected and stored in a GIS at various "working scales" or levels of resolution.
- **Standard for Quality & Reliability**. As digital data becomes available for potential use by a wider group, a standard set of guidelines or criteria will be necessary to determine relative reliability and "fitness for use" by outside organizations with different mandates and information requirements.
- **Standard for Database Documentation.** In an environment where the Provincial Government wishes to reduce the amount of redundant data collection, verification and maintenance and wants to increase the integration potential of different databases, consistent database documentation is an important early step.
- **Standard for GIS Data Interchange.** Potential data incompatibility is already causing concern in New Brunswick, where a number of GIS and CAD systems are in place. Three different interchange formats are currently in use to move data between different organizations. Designating a standard format would reduce the amount of development work required in line departments and reduce confusion among outside producers and users of data.

IMPLEMENTATION CONSIDERATIONS

Standards — like comprehensive base mapping — are a tough concept to sell. The front-end costs in experienced manpower and sophisticated equipment are high, while the benefits (however substantial) accrue at a later date and are difficult to isolate using current methods of departmental cost-accounting.

Dictating the general use of standards will not produce the desired results in the long term. Rather, the only proven incentive to ensuring a standard is adopted is to have one in place. Keys to implementation of any standards include:

- RELEVANCE: There has to be a genuine need (immediate or impending) for the standard.
- EXISTENCE: Database standards cannot be employed if existing data does not cover in the area of interest. If a particular base mapping series or property database is to be designated as a provincial standard, then full geographical coverage should *at least* be available in the areas of greatest use. In addition, the province should commit itself to complete provincial coverage as quickly as possible.
- SUPPORT: Standards should be easy to access and use. Detailed, understandable documentation explaining the contents and characteristics of the database and/or specification should be prepared and kept up-to-date. If possible, meaningful default models and examples should accompany the documentation. To avoid confusion, there should be one definitive source of technical support for the standards in the provincial government.
- FLEXIBILITY: While common standards are desirable for the exchange of data between different groups of users, they need not be imposed in all situations. An agency should not necessarily have to use Provincial standards when exchanging information with another organization sharing the same internal standards for data classification and operating on a similar system. *However,* the agency should be able to relate or transform its information to the common standards when exchanging digital data with an unlike system or when distributing information to the general public.

Rather than developing standards from first principles, every effort must be made to make use of standards implementation and development models from other jurisdictions. Wherever possible, standards development and implementation must take into account the expertise in user departments which have already implemented land and water information systems.

RECOMMENDATIONS

Based on the findings of this project, the following recommendations are made (in order of proposed implementation):

- **Establish the organizational framework.**
  - The Geographic Information Corporation has now been given responsibility for development and implementation of selected land- and water-information standards in New Brunswick.
  - Organize an inter-departmental "Standards Advisory Committee" to ensure input from experienced line departments during standards implementation and refinement.

- Establish limited-term working groups and retain outside assistance to provide expert technical advice on specific standards.

- **Establish interim guidelines during the development period.**

  While government-wide standards are under development, specify that *de facto* standards be well-documented and maintained by the agency most closely connected to the database in question. If cooperative updating is agreed upon as a goal, it should be specified that any return of updated data, or transfers of the same data to a third party, should be carried out in the same format as the original file.

- **Develop standard terminology.**

  This is necessary to make sure that subsequent planning and implementation is carried out in an environment where all contributors share a common understanding of important terms. Rather than starting "from scratch", material from glossaries of GIS- and LIS-related terms already in place in other jurisdictions can be modified to meet provincial requirements.

- **Define a Spatial Referencing Standard.**

  The Province has now adopted the LRIS Base Map as the standard for base mapping.

  - Those departments now considering the use of digital federal 1:50,000 mapping must be aware of the transformations necessary to relate information tied to federal maps to the LRIS base.
  - Agencies like the Emergency Measures Organization which deal with both federal and provincial government agencies will have to deal with potential confusion over different geo-referencing standards as they arise.

- **Standardize parcel definition.**

  The province has now adopted the LRIS Parcel Identifier (PID) number as the standard parcel ID number to be used in all Provincial property-related information systems. This also implies that the parcels delineated on LRIS property maps constitute a standard in their own right.

  - Provincial departments currently maintaining property-related databases will have to ensure that they can relate their parcels to the LRIS standard on a one-to-one, one-to many, or many-to-one basis.
  - If not already included, a field containing the LRIS PID number(s) should be added as a cross-reference to each property record in departmental property databases.

- **Arrange for further investigations before selecting a standard format for GIS data interchange.**

  Detailed testing of different alternatives is recommended before recognizing a single interchange format as the provincial standard. This testing could be carried out by either internal staff or through contract. In either case, the investigation should employ the expertise and facilities already in place at CanLab-INSPIRE at UNB.

- **Develop a standard for database documentation.**

  Standardized documentation will provide interested users with enough information to evaluate the fitness of data for a particular use.

- **Prepare more detailed descriptions of the different databases now in use or under development in various provincial departments.**
  Require all departments to prepare database documentation based on the Provincial standard adopted above.

At this point, the "foundation sub-system" has been standardized and a communication standard has been adopted to facilitate communication between theme databases. While this is still far from the transparent on-line network desired, significant improvements to the current situation will have taken place.

- **Develop a Standard Framework for Data Classification.**
  Since no single classification breakdown emerged as a "standard" from the site interview process, the development and adoption of a standard will depend on whether or not a common framework can be developed which accommodates and categorizes the various types of data maintained in different databases.

- **Standardize Object Representation.**
  A robust standard for object representation is difficult to implement without having a clear picture of potential GIS applications in their respective organizations. Working from the classification framework defined above, a number of alternative representations must be developed for specific feature classifications within a given scale range. Initially, the LRIS specifications could be employed as the interim standard until there is more feedback on the strengths and limitations of that particular series with respect to most applications.

If standards are to be adopted by prospective database developers, then they must be supported by clear documentation and/or comprehensive coverage in the user's areas of interest. To familiarize users with the standards and to refine the standards themselves, all future database and GIS pilot projects should employ adopted standards. If potential conflicts are identified, part of the pilot project should involve developing procedures to relate departmental requirements and specifications to Provincial standards.

Experienced GIS personnel from all agencies must be invited to participate in the standards development process. In organizations where databases are already loaded and information exchange with outside groups is underway, there must be agreement on circumstances where existing standards can still be employed. Further, the ramifications of introducing new standards into these organizations must be defined and allowed for in subsequent budgeting.

TABLE OF CONTENTS

	Page
Executive Summary	i
Table of Contents	viii
1.0 INTRODUCTION	1
1.1 Terms of Reference	1
1.2 Project Methodology	2
1.3 A Working Definition of Standards	3
1.4 Foundation vs. Thematic Layers in a Land Information System	4
1.5 The Rationale for Digital Information Standards in New Brunswick	5
2.0 LAND AND WATER INFORMATION STANDARDS IN NEW BRUNSWICK: THE CURRENT SITUATION	7
2.1 Overall Assessment of Status Quo	7
2.2 Observations from the Site Interview Process	9
2.2.1 Database Documentation	9
2.2.2 Updating of Databases	10
2.2.3 Linguistic Considerations	11
2.2.4 Registry and Land Titles	11
2.2.5 Municipal Issues	12
3.0. GIS DATA INTERCHANGE STANDARDS	14
3.1 Background	14
3.2 Coordinates	15
3.3 Map Projections	16
3.4 Spatial Objects	16
3.4.1 Geometric Objects	16
3.4.2 Cartographic Objects	17
3.4.3 Raster Objects	18
3.5 Object Sharing and Aggregation	18
3.6 Topology	18
3.7 Attributes	19
3.7.1 Feature Code	19
3.7.2 Data Quality	20
3.8 Coding Method	20
3.9 Summary of Interchange Formats	21

Page

4.0 DEFINITION AND ASSESSMENT OF EXISTING STANDARDS 24

4.1 A Standard for Relevant Terminology 24
4.2 Standards for Spatial Referencing 25
4.3 Standard for Parcel Definition & Identification 28
4.4 Standards for Feature or Entity Classification 30
4.5 Standards for Object Representation and Symbology 32
4.6 Standards for Quality & Reliability 34
4.7 Standards for Database Documentation 36
4.8 Standards for Interchange of Land and Water Information Between Different GIS 37

5.0 STANDARDS IMPLEMENTATION CONSIDERATIONS 38

5.1 A "Standards Model" for Land & Water Information 38
5.2 Implementation Philosophy 40
5.3 Proposed Framework 41
5.4 Recommendations 43
5.4.1 Establish the Organizational Framework 43
5.4.2 Establish Interim Guidelines During the Standards Development Period 43
5.4.3 Develop Standard Terminology 44
5.4.4 Define a Standard for Spatial Referencing and Base Mapping 44
5.4.5 Agree on a Standard Definition and Identifier for Parcels 45
5.4.6 Select a Standard Interchange Format for GIS Data Transfer 46
5.4.7 Develop a Standard Format for Database Documentation 46
5.4.8 Prepare Database Descriptions 47
5.4.9 Develop a Standard Framework for Data Classification 47
5.4.10 Standardize Object Representation 48
5.5 Balancing the Needs of Experienced and Non-Experienced Users 48

Selected References 50

Appendix I List of Agencies Interviewed 51

SUPPORTING DOCUMENTATION IN SEPARATE VOLUME

Appendix II Detailed Descriptions of Interchange Formats 55
Appendix III Examples of Database Documentation 80

FINAL REPORT

LAND AND WATER INFORMATION STANDARDS PROJECT
PROVINCE OF NEW BRUNSWICK

1. INTRODUCTION

This report presents the findings of a project that investigated land and water information standards in New Brunswick. It is an extension of efforts to develop a land information network for the Province, as reported in the Final Report of the Landnet Program, Phase 1 (1988). The project methodology and terms of reference will be reviewed and a definition of standards will be presented. The rationale for standards will then be developed, followed by a discussion of the current use of standards in New Brunswick. Finally, an ideal "end vision" for data exchange will be described and recommendations will be made to facilitate movement toward this end vision.

1.1 Terms of Reference

In June 1988, the New Brunswick Surveys and Mapping Committee commissioned The Cabot Group Inc. to investigate standards for land and water information maintained by the Province of New Brunswick and its agencies. Specific objectives of the project included:

- Developing a framework within which existing standards for recording, classifying, coding and storing data, related to land and water management in New Brunswick, can be examined.
- Preparing an inventory of existing standards based on the above framework.
- Identifying any conflicts, inconsistencies or major omissions in current standards which should be addressed by the Provincial Government.

Deliverable items associated with this project include:

- A *Project Report* (this document) which presents the findings of the study and discusses standards issues as they relate to New Brunswick.
- A *Manual of Standards* (delivered under separate cover) documenting specifications for all datasets commonly used by more than one department. The Manual's "Table of Contents" will indicate existing standards as well as standards which will need to be developed or adopted at a later date.
- A *Data Summary Spreadsheet* summarizing both the nature of land and water-related information maintained by agencies, contacted during site interviews, and specific database characteristics.

The longer-term goal of these standards-related investigations is to identify, adopt and/or develop standards, procedures and policies which would facilitate the transparent use of selected datasets among different government departments. It was therefore agreed that the "standards framework" should be *implementation-driven* rather than simply a tool to examine existing standards.

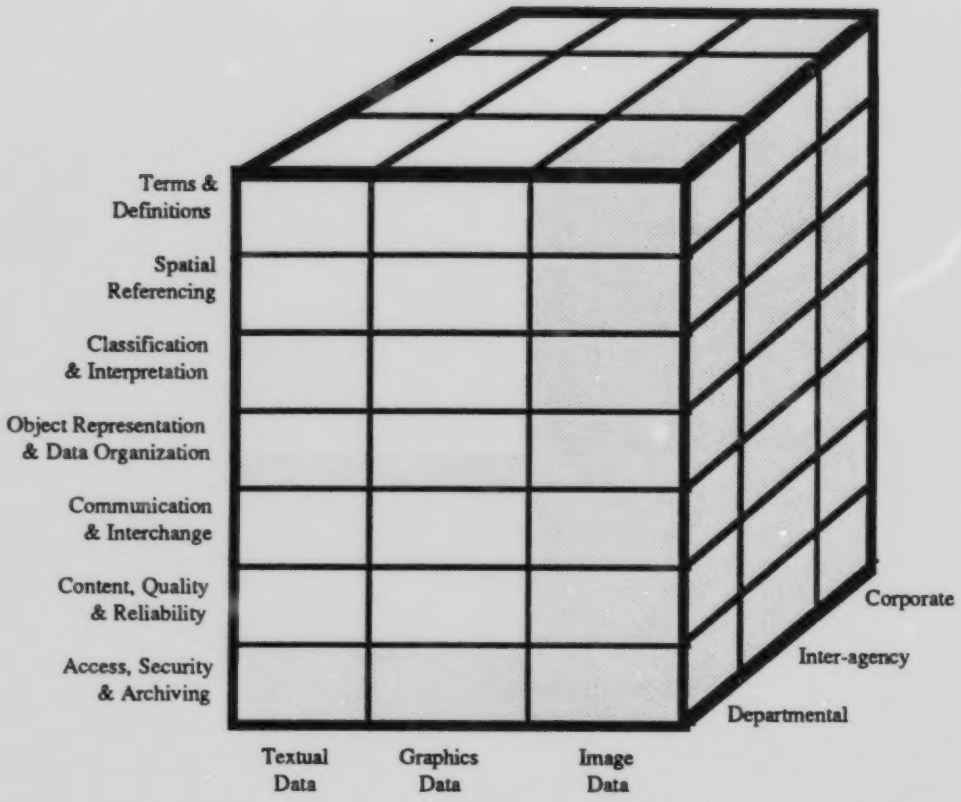

FIGURE 1.

A STANDARDS FRAMEWORK

1.2 Project Methodology

Standards were seen as being *multi-dimensional* in nature: What types of data require standards? What agencies are concerned with standards? What types of standards are possible? To help answer these questions a modification to the approach used by the National Committee for Digital Cartographic Data Standards in the United States was used and is presented in Figure 1. This framework provided a starting point from which to begin a detailed examination of standards and allowed the study to focus on specific issues. More importantly, the framework was used in site interview planning to determine who to interview and the thrust of the interview process. The following types of organizations were contacted for interviews:

- Provincial departments now maintaining land-related information in digital form or expecting to move to digital databases in the next 12-24 months.

- Federal agencies which exchange land-related information with Provincial agencies, municipalities and/or industry.

- Selected departments within a representative municipality and regional planning commission which interact routinely with Provincial agencies.

- Private forestry companies in New Brunswick now using GIS technology.

The final list of those agencies selected for interviewing is found in Appendix I.

1.3 A Working Definition of Standards

Although the term 'standard' is in common use, it is not always used in the same context. In order to reduce the initial confusion over what does and does not constitute a "standard" or standards-related activity, a working definition was prepared and submitted to the Project Steering Committee for approval.

For the purposes of this project, a "standard" is defined *as a specification which describes a dataset (conventional or digital) commonly employed by one or more user groups outside the agency or department which produces that dataset.*. To identify the standards now in use in the Province, over 40 site interviews were arranged with various government agencies, municipalities and private companies in the Province. Based on the interviews, relevant "standards" under the above definition may include, for example,:

- Department of Transportation (DOT) specifications for road classification.
- ATS'77, NAD'27 and NAD'83 reference ellipsoids for basic definition of latitude and longitude.
- Universal Transverse Mercator (UTM) and Stereographic map projections for spatial referencing of point information.
- "Minimum Survey Standards" for legal surveys as developed by the Association of New Brunswick Land Surveyors.
- Canadian Council on Surveying and Mapping (CCSM) feature code specifications for classifying and coding digital map features.
- Land Registration and Information Service (LRIS), Energy, Mines and Resources (EMR), Department of Natural Resources (DNR), DOT and the Provincial Map Book specifications for conventional base mapping.
- LRIS, EMR and DNR specifications for digital base mapping.
- LRIS, DOT, Statistics Canada and municipal maps for municipal street network information.
- LRIS, Provincial Properties Data Center (PPDC) and Department of Municipal Affairs and Environment (DMAE), Assessment Branch parcel definitions and their corresponding identifier numbers.

1.4 Foundation vs. Thematic Layers in a Land Information System

Before geographic information systems (GIS) were introduced, government departments in New Brunswick often created their own special-purpose overlays using topographic base maps and property maps from DNR, EMR and LRIS. By virtue of the common base, users were able to relate different overlays (or themes) to one another for site-specific projects. For example, agricultural planners could examine relationships between soil type, terrain slope and climate. Highway engineers could optimize final route location of proposed roads by taking grade, parcel ownership and presence of existing structures into account.

Since the late 1960's, the concept of an integrated land information system that relates different types of land information to a common base has been widely accepted as the model for most multi-disciplinary land information programs. These programs have typically involved the establishment of: (1) a coordinate reference framework; (2) a series of consistent base maps tied to this framework; and (3) a cadastral overlay delineating parcel boundaries on base map features.

These three components comprise a "foundation sub-system" and are maintained separately, while other land and water information themes are added as layers upon this foundation subsystem (See Figure 2). While the exact content of base maps may be jurisdiction-dependent, as a minimum they would include roads, watercourses, major buildings, transmission lines, designated areas, and other features depending on map scale. To some users, property boundaries may not appear to be a necessary component of the foundation sub-system and simply represent another overlay theme. However, land use and property attribute information maintained in textual form by different departments can be spatially-referenced using parcel identifier numbers contained on property maps. Therefore, property mapping has become a theme frequently required by different users and merits inclusion in the foundation sub-system.

Much of the data collected and mapped by different departments is specialized in nature; e.g., mineral deposits, forest cover, interpolated soil samples, well locations and underground storage facilities. At the same time, most of these same departments also share a common need for information from base maps, property maps and property ownership records. In fact, some users update or correct such information (for their own purposes, at least) when these records are found to be out-of-date or erroneous. However, because of the different standards employed in this process, other users often cannot make direct use of this updated information and must revise these same records themselves as the need arises. As well, in collecting information from different departments, members of the general public must often deal with information portrayed to several different standards of symbology, reliability and timeliness.

This concept of a foundation sub-system with overlying themes is key to the understanding of standards issues as developed in this report. Because all themes are linked to the key subset, linking between themes then becomes possible.

1.5 The Rationale for Digital Information Standards in New Brunswick

While standards were desirable when using conventional hardcopy maps, they were not essential since relatively little *exchange* of data took place between map users. Where some overlap of desired information may have existed, most departments continued to collect and maintain their own map information to internal specifications. For the most part, this alternative was chosen because maps maintained by "outside" agencies — judged by *internal* standards — were either: (a) incomplete; (b) lacking in detail; (c) not current enough;(d) related to an incompatible base map; (e) not positionally accurate enough; or (f) prematurely aggregated using different land-unit definitions.

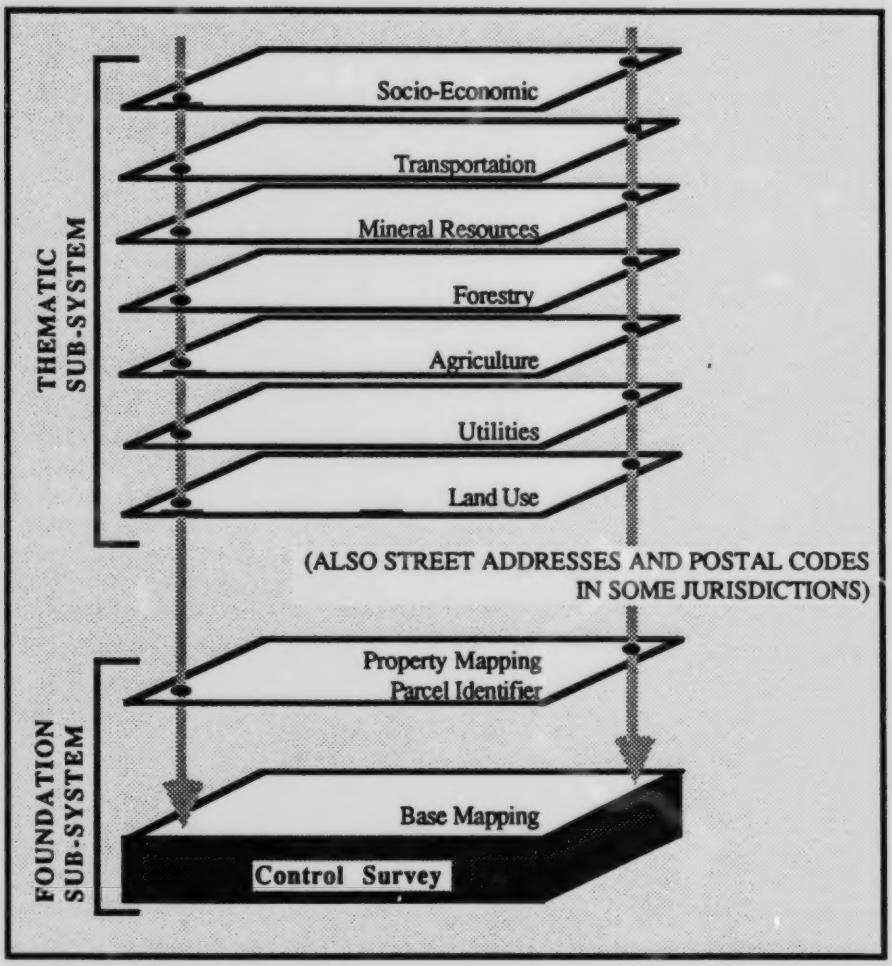

FIGURE 2

THE OVERLAY MODEL FOR PROVINCIAL LAND INFORMATION MANAGEMENT

Although some agencies have automated their land records information, relatively few have implemented full GIS's. Since system acquisition, applications development, and the eventual loading of individual databases all represent a substantial investment to the Province, the potential incompatibility of data collected through different programs cannot be ignored. As more and more Provincial departments in New Brunswick adopt GIS technology there is increasing concern with:

- **Database Loading Costs:** GIS data collection, conversion and database maintenance activities in other jurisdictions have required vast amounts of manpower and computer time and have proven to be very expensive over the life of a program.
- **Availability of Data:** Up-to-date digital data sets containing property and land cover information in particular portions of the Province will be required by more than one user group.
- **Data Translation Costs:** Unlike hardcopy maps and records, the cost to reconcile outside digital data files to a different base, edit and prepare outside datasets for cartographic display and/or geographic analysis, and merge data with internal information can be extremely expensive and time consuming.

In 1988, the Landnet I Project Report discussed the feasibility of establishing a land information network which would link the major individual land related databases maintained by Provincial agencies. More recently, a proposal by NBTel for a pilot project in the Moncton area has also examined the integration of different Provincial datasets within a distributed network. To make full use of data communicated across such land information networks, there must be some agreement on how features are to be defined, coded and represented. In addition, outside users must have enough information about a particular dataset to judge whether or not it is suitable for their purposes.

The development of standards is seen as a means of alleviating some of these concerns. Standards for spatial referencing, classification and communication help establish a "common ground" that can:

- Allow technical and economic efficiencies related to the use of digital data.
- Facilitate communication between different theme overlays that are linked to a common base.
- Assist government departments in consistently defining specifications for digital data submissions from private contractors, federal agencies and other outside groups.
- Enable the Provincial Government to establish more consistent information products for distribution to the general public

Supplemented by the appropriate political and institutional mechanisms, standards can provide the technical basis for limiting the amount of duplication in collection, identification, verification and maintenance of land and water information in New Brunswick. If implemented effectively standards are a means of protecting and even optimizing the Province's investment in digital data.

2.0 Land and Water Information Standards in New Brunswick: The Current Situation

The following sections summarize the current situation with respect to land and water information standards in New Brunswick based on responses obtained during the site interview process. In addition to an overall assessment of the current use of standards, this chapter will also present specific points from the site interviews and discuss some of the strengths and weaknesses of selected standards now in use.

With a few notable exceptions, this report will deal primarily with standards affecting land and water information contained in the "foundation sub-system" described in Section 1.4 and only characteristics of those databases in whole- or partial-digital form will be discussed.

2.1 Overall Assessment of Status Quo

New Brunswick is relatively unique among North American jurisdictions in that so many of its mapping and land records systems (automated or manual) offer complete coverage over the entire Province. Conventional base mapping and property mapping have been complete since 1985 and the "first wave" of digital mapping and GIS technology in LRIS and forestry sectors is now well underway.

Table 1 outlines some of the standards now in use in the Province based on the standards framework presented in Section 1.2. These standards are discussed and compared in greater detail in Section 4.0.

As Table 1 indicates, there are relatively few basic "standards" for spatial referencing competing for use in the Province. The continued use of different standards for hardcopy base mapping (e.g., EMR 1;50,000 vs. LRIS 1:10,000 mapping) stem from traditional map-scale requirements, the need (or lack of need) for Federal-Provincial reporting compatibility, and/or simply the lack of appropriate map coverage in the area of interest at a particular time. These same reasons help explain the creation and subsequent maintenance of graphical and textual property information to different standards by various departments.

Although users have a number of options in terms of standards for spatial referencing, there is little evidence of any common standards for data classification. While not regarded as a problem today, this may constrain communication and use of information in the future, and have serious implications as networks develop.

Using the working definition of standards described earlier, existing standards in New Brunswick can be characterized as belonging to either of two categories of spatially-related information "standards":

- **Category 1:** specifications describing datasets initially planned to become multi-disciplinary standards (e.g. LRIS base mapping and Parcel Index);

Table 1
Standards Currently Employed in New Brunswick

	Textual Data	Graphical Data
RELEVANT TERMINOLOGY	None in Place	None in Place
SPATIAL REFERENCING	N/A	FEDERAL NAD'27 / UTM LRIS 1 ATS'77 / Stereographic (Conventional) LRIS 2 ATS'77 / Stereographic (Digital)
PARCEL DEFINITION & IDENTIFICATION	LRIS PID Assessment Number Prov. Properties (PPDC) ID # Farmland ID # Timber Management Branch	LRIS Prov. Properties Data Centre Timber Management Branch
DATA CLASSIFICATION	Assessment Agriculture Community Plng. DNR Timber Mgt. Branch D.O.T.	LRIS DNR Timber Mgt. Branch D.O.T. Agriculture CLDS
OBJECT REPRESENTATION	N/A	LRIS DNR Timber Mgt.
DATA COMMUNICATION & INTERCHANGE	ASCII Text Files; Application - Specific Translators	DLG ISIF DXF
DATABASE CONTENT, QUALITY & RELIABILITY	TMB LRIS Informal Mechanisms for Independent Verification of Parcel Attributes	LRIS TMB Undocumented Professional Stds. for Interpretation Informal Mechanisms for Independent Verification of Parcel Boundaries

- **Category 2:** the output characteristics of databases which, by virtue of some characteristic lacking in other datasets, became popularly used by two or more departments (e.g., DNR 20-chain mapping or the Property Assessment and Taxation System (PATS) database).

Category 1 includes standards developed with the specific intention of creating a consistent, Province-wide (or even nation-wide) information base for use by a variety of public and private users. With this in mind, detailed documentation concerning content, lineage of information, and data structure of the databases involved is usually available. Examples of such standards include LRIS and EMR base mapping and the LRIS Parcel Index.

Standards included in Category 2 were designed first and foremost as support tools for one departmental program or another. However, because information contained in these databases was also deemed valuable for other programs, subsequent use of this information by other departments for those programs has made the database — by default — a *de facto standard* in the Province, whether or not the originating department views it as such.

In most cases, these Category 2 standards became popular simply because the information contained in them was found to be reasonably up-to-date and accurate. More important, however, *the database coverage spanned most , if not all of the Province and the information contained therein could be readily accessed by outside users*. In fact, they often became de-facto standards because databases that were originally designed as standards were not complete or were not easily accessible at the time the information was required by an outside department. One of the best examples of such a standard is the PATS Provincial Assessment database and, although less obvious, the graphics files of the Timber Management Branch (TMB) GIS and the method of road classification employed by DOT.

Two significant side-issues are connected with Category 2 standards. Specifically:

- because these databases were designed primarily for internal use, information content, quality and structure may change as program requirements and funding levels evolve; and

- since they were not designed for general use, database documentation and specifications may not necessarily be as extensive as Category 1 standards and funding levels may not be in place to upgrade and maintain the documentation to satisfy the demands of a more diverse group of users.

2.2 Observations from the Site Interview Process

The following section highlights specific observations and recurring themes that emerged during the site interview process.

2.2.1 *DATABASE DOCUMENTATION*

With few exceptions, agencies which developed their information systems in-house and now use them exclusively for internal purposes, generally could not provide detailed documentation covering data collection, maintenance and reliability specifications. Database content, on the other hand, could usually be obtained through a review of data

fields named in field listings. In some cases, where records have been maintained by a single individual, that person represents the only reliable source of information about a database. In other cases, where data has been collected over a longer period by different people, standards for reliability, accuracy and content may vary between different "pockets" of the database and/or from project to project.

Where agencies have sub-contracted for data collection services or exchange information routinely with other organizations, there is a greater likelihood of detailed, up-to-date documentation, usually in the form of contract specifications. Even in cases of well-documented graphical databases however, there is often a reluctance to publish any indication of positional accuracy or attribute reliability.

2.2.2 *UPDATING OF DATABASES*

The role of standards in updating databases became a recurring issue during the site interview process. In some cases, databases are initially being created using the best information available to the agency staff at the time of compilation. Since the emphasis is usually placed on obtaining Province-wide (or project-wide) coverage as efficiently as possible, the resulting reliability of some databases may be relatively low during the early stages. These databases are subsequently updated with more timely and reliable information, but newer information is usually only available in "pockets" of the database. Not surprisingly, then, standards within a single database may vary by data type, by geographic area, and over time. Especially when no documentation is available, outside users may assume that all data contained in a particular database has been collected and maintained to the standards currently employed.

UPDATING OF PARCEL-RELATED ATTRIBUTES

In most databases examined, basic ownership, address and parcel information is collected through program application forms and is maintained on an exception-basis only. A few systems —the Farm Land Identification Program database, for example — attempt to verify information supplied by the applicant using data contained in the LRIS or PATS databases. However, given equipment and personnel constraints in most departments, this is the exception rather than the rule. In most of the databases examined, errors or outdated information would be identified and corrected only if errors were found to affect overall program administration. Otherwise, the onus remains with the applicant to provide updated information as necessary.

UPDATING OF MAPPING DATABASES

While specific source agencies may be given the mandate to initially collect one particular type of information (e.g., LRIS or EMR with base mapping), they are not always in the best position to keep this information up-to-date. Other agencies often identify and make note of corrections and updates to these databases as part of routine project or program-related activities. However, because of different procedures and criteria for collecting, locating, classifying and representing data, updates are not necessarily used by the source agency. At best, corrections or updates are incorporated after internally verifying the problem and correcting it to internal specifications. At worst, the source dataset is updated only during the next round of overall coverage which, in the case of base mapping, may be years after the problem has been identified.

Land cover, water, buildings and infrastructure information on LRIS hardcopy base mapping is updated and enhanced independently on a site-specific basis by forestry companies, exploration firms, municipalities, utilities, and the DOT, DNR, DMAE and Department of Agriculture (DOA). However, because these updates are carried out in conjunction with the collection of internal project or program data, features are collected and represented to internal specifications for positional accuracy, classification and symbology. The resulting updates are generally considered to be part of the "theme overlay" rather than part of the base mapping itself. Agencies interviewed were reluctant to even guarantee the consistency of these updates from project to project or site to site, except to point out that the information served the purpose for which it was collected at the time.

Throughout the site interviews numerous examples emerged of cooperative, *ad hoc* efforts to inform LRIS and the Assessment Branch of any errors in either LRIS property mapping, LRIS Parcel Index, or the PATS database. However, the speed and efficiency of subsequent verification and correction efforts depend on the respective specifications used by the parties in question and the type of data under examination. In some cases there has been no clear guidance given to agencies as to who has responsibility for updating.

There is at least one example of an updating procedure in place whereby one agency sets a standard and other agencies using that data are required to update the data to specifications or pay the owner agency to have the updating done. In this case, NBTel made a cooperative updating agreement with NB Power regarding the installation of its fibre-optics transmission line around the Province. One of the conditions for joint use of NB Power rights-of-way was the agreement by NBTel to update — or pay for the updating of — the corresponding right-of-way plans in accordance with NB Power *drafting* specifications.

2.2.3 *LINGUISTIC CONSIDERATIONS*

French and English legends appear on government maps from at least three different departments, although bilingual labels only appear consistently on LRIS mapping. In addition, the Department of Transportation distributes both French and English versions for some of its non-registered information plans. In the Department of Municipal Affairs and Environment, while by-law maps are prepared in both official languages, routine land use maps are prepared in either French or English depending on the predominant language spoken in the project area. Individual practices notwithstanding, the Project Team could find no uniform policy or standard for the translation of maps, plans and land records.

In the longer term, much more effort will have to be expended if full linguistic transparency is to be achieved. Specifically:

- While hardcopy LRIS cartographic products contain bilingual labels, feature codes are expressed in English only. Although queries at this point are still accomplished using feature code numbers, translation tables showing the French-language equivalents should be developed over the longer term.
- Part of the Land Titles Pilot Project in Albert County involves the development of a table of bilingual address equivalents to accommodate queries from both francophone and anglophone users. Depending on the results of this effort, consideration will have to be given to providing a front-end *linguistic interface* if database queries are to be accomplished in both official languages.

Given the active use of both French and English in New Brunswick, there is a definite need for a closer look at linguistic considerations in provincial land information management. Determining the operational concerns *and legal implications* associated with translating map information and land records from one language is beyond the scope of this Project and merits a separate study in itself.

2.2.4 Registry and Land Titles

Registry Office operations have been the subject of numerous documents and reports produced by the Department of Justice and LRIS over the past 10 years. Clearly, the Registry Offices are the most widely-used sources of property information in the Province and, under the working definition for the project, the Registry Indexes represent a Provincial "standard". However, the information stored in these indexes is not organized in a "standard" fashion from County to County. For example, the names of major property owners (e.g., government departments and major corporations) are not expressed consistently from Office to Office. Government agencies may be designated under a specific department name, as "the Crown", as "N.B. Government", or even under the name of the Department Minister. Some Registries will record company names in one manner [e.g., Irving, J.D. (Woodlands) Ltd.] while another will record it differently [J.D. Irving (Woodlands) Ltd.]. Institutional names can pose similar types of problems. As well, some offices sort their indices alphabetically in a different manner than others (e.g., "Mac's" and "Mc's" separately vs. included in alphabetical order).

Since such information was rarely exchanged between registries, this has not been a problem and did not hinder the use of the information by other agencies. However, the recent development of computer systems to compile consolidated indices *and* the potential demand for Registry data in digital form is highlighting the current need for standards.

Registrars have recognized these problems and, in 1987, the Department of Justice issued a new policy for standardizing indices among registry offices. However, information obtained during the interviews indicated that this policy only served to confirm existing rules rather than develop new rules to resolve the inconsistencies. For example, any directives in this policy that conflicted with existing rules were ignored by the Registry offices which had already begun computerizing their operations. In the fall of 1988, a meeting with the Registrars of Deeds, Justice Department officials and LRIS was called to help resolve the problems and a new "Policy on Indexing Real Property in the New Brunswick Land Registry System" was tabled in January 1989. The effects of this policy are still to be determined at the time of writing. In any event, it would only apply to subsequent transactions.

Given:

- the widespread use of the Registry;
- current plans to begin automating some of the indexes in the Registry Offices after reviewing the results of pilot projects; and
- the pivotal importance of the Registry as a potential major supplier of data to a provincial land information network,

— it is clear that actions must be taken to ensure that the record-keeping procedures and conventions are carried out in accordance with commonly-accepted standards *and* that, in

fact, the information being recorded in the Indexes is being expressed consistently across all the Offices. Unless common standards are agreed upon *and implemented in each office*, any efforts to automate will require expensive custom developments to address the individual conventions used in each office.

2.2.5 Municipal Issues

- Representatives from the Engineering Departments in both Moncton and Fredericton emphasized their dissatisfaction with the priorities and design of the 1:2000 digital urban mapping proposed by the Province and stressed the need for larger-scale digital mapping. In this case, "larger-scale" implies standards of content and positional accuracy more in keeping with the 1"=100' conventional maps now in use. Regardless of the final design of the urban mapping program, it is clear that municipal users will require a clear understanding of the standards for content and accuracy characteristics of the digital product when tying in facilities overlays. This process is now underway through separate efforts by NBTel and the Surveys and Mapping Committee and will not be discussed further in this Report.
- In the City of Fredericton, since existing LRIS mapping did not meet internal standards for accuracy and completeness, municipal drafting staff have continued to revise property lines independently of the LRIS Property Mapping Program. The City of Moncton made use of updated LRIS Property maps, but pointed out that they were usually in a position to receive potential updates before LRIS, through their development approval process.
- LRIS Parcel Index files have been only recently updated using civic address information maintained by the City of Fredericton. This particular process was completed for the first time during the summer of 1988.
- Cities typically provide zoning and building permit information to the Assessment Branch on an annual basis, but the data is maintained to individual standards by each party.
- Municipal engineering departments and the Department of Transportation independently maintain information on the length of Provincially-owned roads passing through a municipality. This information is compared on an annual basis and discrepancies are dealt with on an exception-only basis.

3.0 GIS DATA INTERCHANGE STANDARDS

The issue of potential data incompatibility between different systems has already caused concern in New Brunswick, where a number of different systems are currently in use for map production and GIS activities. Since interchange formats represent the major way of moving digital data between systems — and because inter-system communication is a growing management concern at this time, the following sections present a detailed introduction to GIS interchange formats and a qualitative comparison of the major standard formats now of interest. A more detailed description is also provided in Appendix II.

3.1 Background

The proliferation of geographic information systems has resulted in a diversity of data structures for organizing digital map information. Occasionally there is a need to transfer data between these systems of dissimilar data structures. A system generating data is called a donor, and a receiving system is called the receiver. The purpose of an interchange format is to serve as a neutral data structure through which the donor and receiver can exchange data. A published and widely followed neutral structure is called an *interchange standard.*

Ideally there should be a single standard. However, because standards development is an evolutionary process, this goal is unattainable at the present time. As a result, several interchange "standards" exist and many more are under development. This section will examine the characteristics of the following interchange standards:

- **CCSM** - As proposed by the Canadian Council on Surveying and Mapping.
- **DLG** - The Digital Line Graph Format developed by the United States Geological Survey (USGS) to distribute USGS digital maps.
- **DXF** - The ASCII Drawing Interchange Format designed to exchange data between AutoCad systems, and between AutoCad and other computer aided design (CAD) systems.
- **MACDIF** - The Map and Chart Data Interchange Format developed by several Canadian Federal and Provincial departments to exchange mapping and charting data.
- **SIF** (or **ISIF**) - The Intergraph Standard Interchange Format designed to exchange data between Intergraph systems, and between Intergraph and other CAD systems.

Two of the standards, DXF and SIF, are industrial standards which were originally designed for users of particular systems. In this case they are both related to CAD systems which were adopted for mapping applications. The remaining three are public standards. The two Canadian standards, CCSM and MACDIF, are still under development. While originally intended to distribute USGS data, the American DLG format is now also used by other systems to exchange digital map data.

The reason for selecting these standards for study, among numerous others, is either because of their current use within the New Brunswick, or because of their potential national impact. The following table shows the major computerized mapping systems installed within New Brunswick and the Maritimes and the interchange formats supported by them:

System	*DLG*	*DXF*	*MACDIF*	*SIF*	*Sample of others*
Arc/Info	•	•		•	AMF,DIME,IGES,...
CARIS	•	•	•	•	NTX,...
GeoBased	•	•		•	DIME,...
Intergraph	•	•		•	AMF,DIME,IGES,...
SPANS	•	•		•	DIME,...

No single format has emerged to date as *the* standard in the Maritimes. DLG has been used predominantly by MRMS, LRIS and the College of Geographic Sciences for transferring files between CARIS, GeoBased, Arc/Info and SPANS systems. SIF has been used by Eastcan to transfer CARIS files on to Intergraph systems at ADI Ltd. and NB Power. Finally, DXF has been used to transfer CARIS files to AutoCad systems in NB Tel.

While a more detailed overview of the five standards examined is given in the Appendix II, this section compares the major characteristics of these standards in terms of coordinates, map projections, spatial objects, topology, attributes, and coding method.

Depending upon the application, some of these characteristics are more important in an interchange format than others. Those that are essential in a GIS environment have been shown in **bold** in the comparison tables.

3.2 Coordinates

The majority of data in an interchange file consists of points which are based on a cartesian coordinate system (i.e. expressed in x,y, and sometimes z). The dimension, resolution, and method of encoding coordinates affects the efficiency of an interchange format.

While CCSM and DLG are strictly related to two-dimensional and three-dimensional features, respectively, the other standards allow points to be represented in either dimension.

Resolution refers to the smallest measurement that can be represented. It is directly related to the number of significant digits used to define the coordinates. For example, the largest coordinate that can be used in a SIF file is 4,294,967,295 with 10 significant digits. If such a system is used to map the entire world, which is about 40,000 km long at the Equator, the resolution of each coordinate would be slightly better than one centimetre. MACDIF, on the other hand, allows an unlimited number of significant digits.

Concerning data encoding, all interchange standards examined support the absolute coordinate mode which references a unique origin. The relative mode uses the previous

point as the origin which has the effect of reducing the magnitude of the coordinates if the points are closely spaced. MACDIF goes a step further to also provide differential and incremental modes for further data compaction.

These three characteristics of a coordinate system are considered to be essential to topographic mapping and all of the standards examined meet the minimum requirement.

Standard	***Dimension***	***Resolution***	***Mode***
CCSM	3	15 digits	abs
DLG	2	12 digits	abs
DXF	2 or 3	14+ digits	abs /rel
MACDIF	2 or 3	Unlimited	abs /rel /dif /inc
SIF	2 or 3	10 digits	abs

3.3 Map Projections

The cartesian coordinate system is an artificial one related to the real world by a map projection. Without a map projection, it is impossible to obtain the geographic coordinates (in degrees latitude and longitude) of points in the file, or to convert to another map projection.

The number of map projections supported varies from one standard to another. However, it is felt that the inclusion of more map projections can be accomplished quite easily and is therefore not a major concern. Of the three standards supporting map projections, only two, CCSM and MACDIF allow the direct coding of points in geographic coordinates (i.e., in degrees latitude and longitude). The support of projection parameters in an interchange standard is considered essential.

Standard	***Projection***	*Geographic*
CCSM	yes	yes
DLG	yes	-
DXF	-	-
MACDIF	yes	yes
SIF	-	-

3.4 Spatial Objects

The most basic elements in an interchange file are points. Higher level objects, such as lines and areas, are constructed from points and even more complex objects are formed

from these. Spatial objects comprise three types; geometric, cartographic and raster. The variety of spatial objects supported, and the ability to form compound objects is a measure of the flexibility of the standard.

3.4.1 *GEOMETRIC OBJECTS*

Two regular geometric objects often used in mapping are circles and ellipses with circular arcs and elliptic arcs being special cases of each. These regular objects can always be represented by a series of points. However, a more compact and precise representation is to use geometric parameters. For instance, a circle can be uniquely defined by a centre and a radius. The support of geometric objects is not essential for an interchange standard since they can always be represented by a series of discrete points.

Standard	*Circle*	*Ellipse*
CCSM	yes	yes
DLG	-	-
DXF	yes	yes
MACDIF	yes	-
SIF	yes	yes

3.4.2 *CARTOGRAPHIC OBJECTS*

There are two cartographic objects that are so distinctive they deserve special attention. One is curved text which is a text string following a curved feature such as a river. If a standard fails to preserve curved text, additional editing must be performed by the receiver. It is therefore considered to be an essential characteristic of an interchange standard.

The other cartographic object of concern deals with the placement of height data when representing contours. DLG uses a two-dimensional coordinate system and requires the height to be part of the feature code. DXF and MACDIF allow a single z coordinate to be associated with a two-dimensional line. CCSM, which is based on a three-dimensional coordinate system, is unable to handle contours efficiently because the redundant z coordinate cannot be removed from the data.

Standard	*Curved Text*	*Contour*
CCSM	-	-
DLG	-	Z in feature code
DXF	-	yes
MACDIF	yes	yes
SIF	-	-

3.4.3 *RASTER OBJECTS*

None of the standards studied currently support objects in raster form. Raster objects are included in MACDIF documentation but they are not defined at this time. The Canadian Centre for Remote Sensing (CCRS) is currently conducting studies in the interchange of raster objects. Since more and more information will be expressed i raster form in the future, developments in this area should be followed closely.

3.5 Object Sharing and Aggregation

Object sharing concerns the ability of two different features to share the same location without duplicating coordinates. For instance, a feature called 'river' and another called 'parcel boundary' might coincide in space because the boundary follows the river. It is highly desirable that an interchange standard allow objects to share coordinates. The most common way of achieving this is to allow multiple feature codes.

Object aggregation allows the grouping of several objects to form either a new object or a separate layer of information. For mapping, the main interest is to aggregate features into different thematic layers.

Neither of these two characteristics are considered essential to an interchange standard for topographic mapping.

Standard	*Sharing*	*Aggregation*
CCSM	yes	-
DLG	yes	-
DXF	-	yes
MACDIF	yes	yes
SIF	yes	yes

3.6 Topology

The position, shape, and size of spatial objects is only one aspect of geometry. The other aspect involves topological relationships among the objects, which are also preserved in a geographic information system and should therefore be supported by an interchange standard. Some examples of these relationships include:

- The enclosure relationship between a polygon and its islands.

- The connectivity relationship between lines in a network.

A loop is a classic way of encoding polygon topology and is useful for some applications. Although DXF, SIF and MACDIF allow a line to be closed to form a loop, DXF does not

recognize islands. A more compact and structured method is based on lines that start and stop at intersections. Networks (such as highway systems), and polygons are then constructed from these special lines. For flexibility an interchange standard for geographic information systems should support this structured method.

Topology is not considered essential for topographic mapping. Even for geographic information systems, topology can usually be reconstructed by the receiver, although it is often very time consuming.

Standard	*Closed Loop*	*Structured*
CCSM	-	yes
DLG	-	yes
DXF	no holes	-
MACDIF	with holes	yes
SIF	with holes	-

3.7 Attributes

Attributes are information, such as text or numerics, associated with spatial objects. Since the nature and scope of these attribute values differ from system to system, there is little attempt to standardize them. However, it is gradually being recognized that an interchange standard would allow the transfer of attribute data. To do this a description of the format and its meaning would be contained in the file. Such a description is sometimes called a data dictionary.

Two elements of attribute information, the feature code and data quality descriptions, have been singled out because of their significance to mapping. The feature code is an identification for the feature type. For instance, all paved highways could belong to a certain type and be given the character code 'PAVED-HIGHWAY' or the numerical code '1201', according to the convention adopted by the mapping organization. Data quality is related to the source, precision, and completeness of data.

3.7.1 *FEATURE CODE*

Feature codes have become so popular that there are standards (such as the one proposed by CCSM) for feature coding. Whether a feature code is in character or numeric form is not extremely important although those in character form allow for more variation. However, the length of a feature code is significant because it affects the degree to which a feature can be classified. The conversion from a system of long codes to that of a shorter codes would result in truncation and hence the loss of information. Feature code length then, is considered to be the only critical attribute characteristic.

Most other attributes are non-standard and vary from one application to another. A proper format for the transfer of attributes should include the data dictionary which would describe what information is contained in the attribute records and how it is stored. Without the data

dictionary, it would be impossible to decode the attribute data unless the donor and receiver made other arrangements, such as an informal description on paper.

Standard	*Feature Code*	*Other Attributes*	*Data Dictionary*
CCSM	12 characters	yes	yes
DLG	7 Digits	-	-
DXF	-	yes	yes
MACDIF	chars. unlimited	yes	under develop.
SIF	-	yes	-

3.7.2 *DATA QUALITY*

The three aspects of data quality addressed here are source, positional accuracy, and completeness. Source describes how, when, and where the data are obtained for a map. Positional accuracy describes, for example, the precision of instruments used in data collection and the scale range within which a map is valid. Completeness describes whether there is missing information that can result from data sampling, generalization, or incomplete survey.

In this comparison, the distinction is made between systems that can assign data quality descriptions to individual features from those which can only assign them globally to the entire map. Data quality descriptions are essential for topographic mapping.

Standard	*Source*	*Positional Accur.*	*Completeness*
CCSM	global	global	global
DLG	global	global	-
DXF	-	-	-
MACDIF	feature	feature	feature
SIF	-	-	-

3.8 Coding Method

In order to carry out an interchange, data must be encoded and physically transmitted from the donor to the receiver. Data transmission can be affected by a number of factors, including the use of magnetic tapes and computer networks. The efficiency of data transmission depends largely on the compactness of the encoded data stream.

Most of the interchange standards examined communicate using ASCII characters because they can be recognized by almost all computers. Moreover, they can be printed in hardcopy form for easy reading. However an ASCII file tends to be bulky.

Binary numbers are more compact. For instance, the number 255 in binary form requires only one third the storage space of the same number in ASCII form. Unfortunately, there

is yet little agreement among computer manufacturers on how binary numbers should be represented thus making the format less desirable for data interchange.

MACDIF uses a scheme based on octets (or 8-bit bytes). This is a variation of the binary method but is more flexible in that it can handle numbers of arbitrary length, and uses a standardized method of representing numbers in binary form.

Whether or not a coding method uses tags affects the degree of data compression. A tag is an identifier placed at the beginning of a record, command, or data item for identification. For example, the description X = 10.3 carries the tag "X =" which identifies what follows as an x coordinate. The advantage of this method is that the absence of a tag implies the absence of an item. Tags, if over used, can consume significant amounts of storage. MACDIF uses a scheme that allows the use of short tags.

An alternate method to using tags, is to imply the meaning of an item by its position. Eliminating tags does save space, but the fact that they are position dependent does not allow for the efficient coding of default values which avoid the repeated specification of a globally applied value. For example when specifying a constant height for a contour line and using a coding method, which is position dependent and allows only encoding of three-dimensional points, all z coordinates must be present and occupy space. This redundancy can be avoided by using tags because the absence of a z tag for the second and subsequent points implies that the height specified for the first point prevails.

The coding method does not affect the quality of the interchange, but files that use short tags are more flexible and compact.

Standard	*Format*	*Tagged*
CCSM	ASCII	-
DLG	ASCII	-
DXF	ASCII	yes
MACDIF	octet	yes
SIF	ASCII/binary	yes

3.9 Summary of Interchange Formats

Five standards have been compared in terms of several important characteristics for spatial data interchange. A summary of the comparison is presented in Table 2, and more detailed background information on each standard is presented in Appendix II. The following table shows the main strength and weakness of each.

Standard	*Main Strength*	*Main Weakness*
CCSM	Object sharing	Always three-dimensional
DLG	Topology	Two-dimensional
DXF	Simple	Long file
MACDIF	Comprehensive	Complex
SIF	Simple	No topology

TABLE 2
SUMMARY OF INTERCHANGE FORMAT COMPARISON
(Characteristics in **bold** are considered to be essential in a GIS environment.

Characteristic	CCSM	DLG	DXF	MACDIF	SIF
COORDINATE SYSTEM					
Dimension	3	2	2 or 3	2 or 3	2 or 3
Resolution	15 digits	12 digits	14+ digits	Unlimited	10 digits
Mode	abs	abs	abs /rel	abs /rel /dif /inc	abs
MAP PROJECTIONS					
Projection	yes	yes	-	yes	-
Geographic	yes	-	-	yes	-
GEOMETRIC OBJECTS					
Circle	yes	-	yes	yes	yes
Ellipse	yes	-	yes	-	yes
CARTOGRAPHIC OBJECTS					
Curved Text	-	-	-	yes	-
Contour	-	Z in feature code	yes	yes	-
OBJECT SHARING AND AGGREGATION					
Sharing	yes	yes	-	yes	yes
Aggregation	-	-	yes	yes	yes
TOPOLOGY					
Closed Loop	-	-	no holes	with holes	with holes
Structured	yes	yes	-	yes	-
ATTRIBUTES					
Feature Code	12 characters	7 digits	-	chars. unlimited	-
Other Attributes	yes	-	yes	yes	yes
Data Dictionary	yes	-	yes	under develop.	-
DATA QUALITY					
Source	global	global	-	feature	-
Position Acc.	global	global	-	feature	-
Completeness	global	-	-	feature	-
CODING METHOD					
Format	ASCII	ASCII	ADCII	octet	ASCII/binary
Tagged	-	-	yes	yes	yes

In general, the two industrial CAD standards, DXF and SIF, are capable of exchanging simple graphic data. Although they support a large number of geometric objects such as surfaces and solids, they support neither topology nor map projections. DXF provides an adequate scheme for interchanging attributes and their data dictionary.

Among the other public mapping standards, MACDIF is the most comprehensive, most compact, and is the only one that supports charting applications. MACDIF, however, is complex and has yet to be generally accepted.

It should be pointed out that this comparison considered only items explicitly coded in the interchange format. For instance, a standard might not provide a place in its format to store map projection parameters, but might support general attribute data so that these parameters can be treated as attributes. This approach requires negotiation (usually resulting in an external document accompanying the data) between the donor and receiver during the interchange. Without such negotiation, the standard itself provides no mechanism for the receiver to locate these "hidden" data.

All interchange standards will evolve over time, although some will change faster than others. Variations (or dialects) of a particular standard will also evolve. The CCSM and MACDIF standards are not yet official, and a new draft of the CCSM standard is due to be released early 1989.

4.0 DEFINITION AND ASSESSMENT OF EXISTING STANDARDS

This section discusses the major classes of standards as identified in Figure 1 and their relevance to land and water information management in New Brunswick. General characteristics and benefits of each standard are defined and the similarities and differences between respective standards now in use in New Brunswick are discussed.

4.1 A Standard for Relevant Terminology

A standard set of relevant terms and definitions for land- and water-related information management is necessary to facilitate better understanding of the issues at stake and to reach consensus on the detailed planning of standards. Such a glossary would contain commonly accepted and understood descriptions of relevant terminology related to existing and proposed spatial information standards, programs and applications in New Brunswick.

COMPONENTS:

- Commonly-used data processing and communications terminology (e.g., UNIX, optical disk, relational database, flat file, ASCII file, multiplexor, open architecture, etc.).
- Generic operations and functions of different geographic information systems used in the Province (e.g., reclassification, buffer/corridor generation, raster/vector conversion, neighbourhood analysis, slope maps, draping, etc.).
- Terms describing characteristics, properties, structures and formats of information contained in different databases (e.g., "network" vs. "polygon" topology, NTX format, DLG, polygon labels, TIN, Area Master File, etc.).
- Selected terminology commonly used in Provincial applications which may be of interest to — or have ramifications for— outside users.

BENEFITS:

- Would develop a "common language" in subsequent technical, procedural and policy-level discussions.
- Would reduce confusion of new users having to deal with commonly-used acronyms and "buzzwords".
- Would describe content, structure and procedures and policies related to information systems and programs in different departments in a more consistent manner using commonly understood definitions.
- Could explain exceptions or enhancements to any common terms or practices using commonly understood and well-defined common terminology.
- User departments, funding agencies, vendors, contractors, and potential outside users would share a basic understanding of relevant terminology during investigations and negotiations.

CANDIDATE STANDARDS:

- None in place.

4.2 Standard for Spatial Referencing:

A complete spatial referencing standard provides the practical basis for explicitly locating, recording and retrieving features on the ground. More importantly, it provides the mathematical basis for consistent digital coverage across the Province for use within GIS applications.

Referencing standards are employed either *explicitly* — where x,y-coordinates of discrete points are surveyed in the field or derived from the map grid for some purpose; or *implicitly* — where features are added or overlays are related to an existing base map by reference to identifiable features rather than explicit coordinates (e.g., road intersections, fence lines, building corners, etc.) The manner in which an organization references its data will determine its attachment to a particular aspect of a spatial referencing standard.

COMPONENTS:

- a commonly-adopted Provincial coordinate framework based on a reference ellipsoid and an agreed-upon map projection, supported by:
  - an accurate network of monumented control points suitably configured to ensure efficient and cost-effective use;
  - full documentation supporting the efficient operations within the network, including data describing point reliability, inter-visibility, etc.; and,
  - full documentation describing the actions required and the implications involved in transforming data from one coordinate system (map projection, reference ellipsoid, etc.) to another.
- a common standard for digital base mapping, supported by:
  - a range of base mapping products designed to satisfy the respective minimum acceptable requirements for content, resolution, accuracy and currency for specific user groups in New Brunswick;
  - complete and consistent digital coverage (horizontal and vertical) at scales commonly employed by users in New Brunswick;
  - full documentation describing the content, structure, date of production and accuracy of the various products which comprise the base mapping standard;
  - detailed documentation, where required, of the procedures required (and potential problems encountered) in: (a) translating the graphics and feature-code data from the original system to the user's system; and (b) transforming mapped information from one referencing standard to another;
  - availability of technical support and assistance to Provincial users.
- a common cross-referencing system for indexing and numbering of maps and plans.

NOTE: Since some systems now posses address-matching capabilities, spatial referencing standards in some jurisdictions also include unique specifications for municipal street addresses and postal codes.

BENEFITS:

- Would reduce the digitizing of the same basic map information by different departments at different times (e.g., roads, water, structures, and relief).
- Would facilitate the collection and cross-referencing of overlay information from different sources.
- Would eliminate the processing required to transform data from one coordinate system to another.
- Would reduce the processing, interpretation and editing effort associated with transforming data from one base mapping standard to another.
- One of the factors in facilitating the cooperative updating of base mapping by many different departments.

CANDIDATE STANDARDS:

- The *Original Federal-NTS Standard* defined by the North American 1927 Datum (NAD'27), the Universal Transverse Mercator (UTM) map projection, and the conventional NTS map sheets prepared at 1: 1 million, 1:500,000, 1:250,000, 1:50,000 and 1:25,000-scales.
  Examples of Databases Using This Standard: Emergency Measures Organization and Water Resources Branch, DMAE; Statistics Canada; Canada Land Data System; N.B. Agriculture Land Information System.
- The *Original LRIS Standard* defined by the 1977 Average Terrestrial System Datum (ATS'77), the N.B. Stereographic Projection, and the conventional LRIS orthophoto and urban map series.
  Examples of Databases Using This Standard: DNR Timber Management Branch; Water Resource Branch, DMAE; N.B. Agriculture Land Information System.
- The *Digital LRIS Standard* defined by the same datum and map projection, but using the detail and features as represented on LRIS digital mapping completed since 1983.
  Examples of Databases Using This Standard: N.B. Agriculture Land Information System; DNR Crown Lands (Aquaculture).
- The *Forestry GIS Standard*, which shares the same basic characteristics as the "Original LRIS Standard" and has the advantage of full topological structuring, but contains less base map content.
  Examples of Databases Using This Standard: DNR Branches dealing with forestry and wildlife; private forestry companies in New Brunswick.

Another combination — the *Digital Federal-NTS Standard,* — is defined by the same datum and map projection as its conventional counterpart but using detail and features as represented on digital 1:50,000 maps. However, it is not yet used regularly in New Brunswick because new digital 1:50,000 NTS coverage is only found in the western part of the Province.

More detailed summaries and specifications for each of the standard databases can be found in the Manual of Standards to be delivered under separate cover. A detailed comparison of

the Digital LRIS and DNR Forestry GIS base mapping was carried out as part of the Landnet Project and a report is available through the Surveys and Mapping Committee.

The following points should be considered by GIS users when dealing with these particular standards:

- Mathematical transformation between the "Original LRIS" and "Original NTS" coordinate systems (and vice-versa) is possible. LRIS in both Summerside and Amherst possess the necessary datum-conversion software and most major system vendors have the software necessary to transform between different map projections.
- Representation of detail on base maps will vary depending on the original compilation scale, date of photography and respective specifications for content, classification and symbology. This can have significant ramifications when specific polygons or areas of interest are related to major streams, rivers and lakes and coastal boundaries. Referencing information to features shown on one base and then transforming it to another base can result in numerous small discrepancies and/or "sliver polygons" which may or may not be a problem, depending on the user's criteria.
- The base maps associated with each of these standards differ in terms of content, classification and structuring. Since it was designed primarily as a base for forest management purposes, the DNR Forestry GIS base contains less base map detail than the other standards. Content differences between the LRIS and NTS base maps are concerned primarily with criteria for classification and symbology.
- Whether or not the graphics files were *topologically cleaned and structured* is an important consideration in cases where the data is likely to be used in subsequent overlay applications. Because it was designed with this in mind, roads and drainage on the DNR Forestry GIS base map are topologically-structured. The basic relationships required to build topology are in place in the Digital LRIS base, but the topology itself has not been built in the files produced to date. Digital NTS maps produced to date have been produced with cartography rather than analysis purposes in mind, and extensive editing has to take place on these files before the topology can be constructed on the user's GIS.

4.3 Standard for Parcel Definition & Identification

The need for agreement on a common parcel identifier has been discussed at Provincial Land Information Conferences for more than six years. Even so, some Provincial users were not aware that there is not always a 1:1 correspondence between the parcels maintained by LRIS, the Assessment Branch of DMAE, and the PPDC. (Specific differences between the three were described in detail in the Landnet Final Report.) While there is a demonstrated need for different "management units" to accommodate Assessment and Crown lands management requirements, among others, these parcels should be at least relatable to the corresponding LRIS property mapping and Parcel Index.

COMPONENTS:

- A common definition for a 'standard parcel' which corresponds to the management parcels maintained by different departments on either a one-to-one, one-to-many or many-to one basis.
- A unique identifier number for each standard parcel.
- Specifications describing the method of collecting, verifying, coding and digitizing the graphics and attribute data associated with each parcel.
- Complete, consistent and up-to-date digital coverage of these standard parcels across the Province capable of being maintained and improved on a transactional basis.
- Documented guidelines for the definition of management parcels with respect to the standard parcel and description of any variations from this definition, where applicable.
- Revisions made to existing database coding to allow access of parcel-related information using the standard parcel identifier.

BENEFITS:

- Rather than having to know department or program-specific identifiers to find parcel-related information in each database, users only have to know one standard identifier for each parcel.
- Agreement on standard parcel definitions would facilitate the cooperative updating and purification of LRIS property mapping using information from other departments maintaining boundary information.

CANDIDATE STANDARDS:

- While the maps and records available in individual departments and private companies may be more accurate within a limited area of interest, the LRIS Property Mapping Series contains the most generally-complete collection of properties across the entire province (Crown and freehold).
- The LRIS PID number (used to link the property mapping to the LRIS Parcel\ Index) is generally used as a cross-reference wherever people make use of the corresponding property maps.

- The Assessment Account number is also widely used — particularly within the provincial government — by organizations interested in obtaining assessment data. This number came into common use because the PATS on-line database was much more convenient to use than the LRIS Parcel Index and, until recently, the only parcel-related access to the database was via the Assessment Account Number.

4.4 Standards for Feature or Entity Classification

Developing standards for the classification of features and entities is one of the most important aspects in the standards debate and is perhaps the most difficult to resolve. Integrating information — or even determining the suitability of information available from various sources — requires a common framework for the classification of thematic information. Working within a "common framework" does not necessarily mean that everyone must use the same terminology, nor does it mean they should collect their land-cover and land-use information to the same level of detail. As shown in Figure 3, it does mean that the various elements, features and characteristics of land and water collected by different departments should be able to be placed on one-or more 'levels' within a common hierarchy.

COMPONENTS:

- A common framework for the hierarchical classification of land and water information (perhaps extended from the CCSM standard or from some other effort now underway).
- An extendable hierarchical coding system which corresponds to the elements within the framework.
- Individual mechanisms and/or translators to relate the different types of data stored in individual systems to the standard on a one-to-one, one-to-many, or many-to-one basis.

BENEFITS:

- Will overcome the need for common feature coding schemes between different departments, programs and systems, which may be difficult and costly to standardize in any case.
- Will help identify cases where the same element is stored and maintained by separate organizations.
- Will differentiate between types of data which may appear to be the same to outsiders but are actually different. "Parcel" information is a good example of where such confusion exists among different departments.
- Users can become aware of the inter-relationships existing between their own databases and those created by other organizations.
- Users would be able to determine the applicability of outside information to their own program and the processes required to relate one to the other.
- Is necessary as an inventory and as a "road map" for the networking activities described in the Landnet Report.
- If accompanied by corresponding feature codes, this framework could provide novice GIS users with a model for the coding of information within their planned systems.

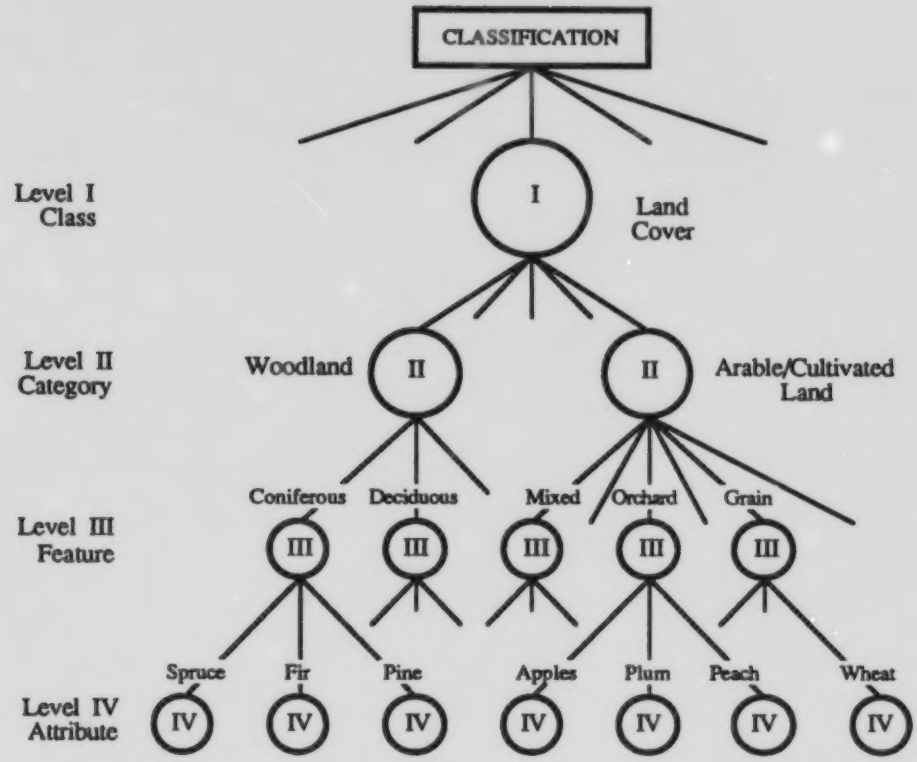

FIGURE 3
EXAMPLE OF CCSM CLASSIFICATION BREAKDOWN

CANDIDATE STANDARDS:

- No overall standards in place at this time.
- The road classification system used by the *N.B. Department of Transportation Highway Book* is also employed by LRIS, the DNR Timber Management Branch, and various forestry companies in their files. However, each one of these groups has extended the framework to some extent, and LRIS was directed to aggregate the road classes somewhat differently than the Drafting Section within DOT. These minor differences should be investigated further to confirm the true "mapped relationships" between different road classes in each system.
- At least six different Provincial organizations collect and maintain various types of land use information on a parcel basis for their respective programs, and a common thread through all the interviews was the potential value of a common framework for land-use classification.
- The hierarchical classification framework developed by the CCSM — now used as the model for LRIS classification and feature coding — was developed after consultation with many federal user departments and might represent an appropriate starting point. However, it would have to be extended to include detailed property and land use classifications.

4.5 Standards for Object Representation and Symbology

'Object representation' is the digital equivalent of cartographic symbolization and, as such, is closely linked to all the other standards mentioned above. Standards for object representation govern how features are to be collected and stored in a GIS at various "working scales" or levels of resolution — e.g; deciding to collect and store only the centerlines of roads beyond a certain scale or, representing certain features as a single point (to be subsequently symbolized). By having a common standard for representation of features in digital form, agencies can use the same digital map file and continue to make use of cartographic symbology that their own staff understand.

When agencies do not share common standards in this area, some re-digitizing or editing of base map data will be required. For example, the DNR Crown Lands Branch made use of LRIS digital map files in a recent aquaculture-related GIS pilot project. However, because roads in the LRIS files are represented as single-line features only, LRIS was contracted to re-digitize all roads to show them as double-line features in the digital file, a representation more familiar form DNR.

While this standard will allow departments to use their own unique symbology internally, a single "default set" of cartographic symbols will be required when data from different sources is accessed on-line by the general public or by departments with no GIS capability of their own. Default sets will be required for both medium- (resource) and large- (urban) scales of mapping.

COMPONENTS:

- A standard set of definitions describing the digital representation of different elements within the classification framework (e.g., single point; single-point plus orientation point; line; polygon, etc.). These definitions may vary depending on the "working scale" or resolution of the digital mapping series in use;
- A default set of cartographic symbols and line-types corresponding to the elements stored in the standard classification framework and developed to be consistent with the object representation standards.

BENEFITS:

- Digital cartographic information collected at a particular level by one organization would be represented in a form consistent with the same data collected for a different application in other organizations.
- Will help ensure that data digitized by one organization is potentially useable by another organization without extensive re-digitizing.
- Avoids the need for Province-wide standards for cartographic symbology, which will vary by necessity due to differences in national, international, or profession-wide standards; BUT provides a default set of symbols for use by groups who have no symbology standards of their own but still require provincial information products (e.g., the general public).

- Focuses greater inter-departmental attention on the data itself and their potential applications, rather than on more traditionally and controversial cartographic issues.

CANDIDATE STANDARDS:

- No standards for object representation formally developed.
- Object representation specifications developed in conjunction with EMR, LRIS and DNR Forestry GIS programs qualify as "default standards" in this regard.
- Cartographic symbology "standards" include specifications for LRIS and EMR digital mapping. However, less emphasis is being placed on delivering cartographic-quality, hardcopy products because experienced users usually have developed their own symbol libraries. Further, the system-dependence of the LRIS and EMR symbol libraries and line codes remains to be determined.

4.6 Standards for Quality & Reliability

Although agencies collect and maintain information that may not necessarily be complete, up-to-date, or positionally accurate to stringent specifications, it nevertheless meets internal requirements within existing time and budgetary constraints. The information suits the purpose for which it was intended. However, as more and more digital data becomes available for potential use by a wider group, a standard set of guidelines or criteria may be necessary to determine relative reliability and "fitness for use" by outside organizations with different mandates and information requirements.

Unlike much of the textual data stored in departmental land and water information systems —which is either correct or incorrect, or contains a statistically estimated error— the reliability of mapped information is usually not as clear. In many cases, the most positionally accurate information is not the most up-to-date. The most up-to-date information may not be suitable in terms of attribute reliability. This situation can be further complicated as updates to the original data, often collected to different specifications again, are incorporated into the database.

COMPONENTS:

- Minimum specifications, categories or classes which quantify data collected in terms of:
  - estimated reliability of attribute information;
  - estimate of data rendered out-of-date in a given period;
  - date of last overall update;
  - updating cycle;
  - positional accuracy (or range of accuracies) of information stored;
  - method of data verification; and
  - other issues as determined by the Provincial Government.

BENEFITS:

- Provided the interpretations are consistent, all organizations receive a consistent initial picture of the relative reliability of an external dataset without compiling an internal assessment of the database.
- Used in conjunction with other standards, these quality and reliability specifications can help occasional or novice users identify the dataset best suited to their requirements.
- Users can get a picture of the reliability of their results when integrating and overlaying thematic information from various sources.

CANDIDATE STANDARDS:

- No overall standards in use by more than one department.
- LRIS digital mapping specifications indicate positional tolerances and date of photography, but updating carried out in line departments is not necessarily incorporated to the same specifications.

- Minimum Standards for Legal Surveys in New Brunswick are mentioned by most departments, but virtually all organizations interviewed indicate that their surveys are conducted to much more stringent specifications. Some agencies also commented that these standards were confusing, since they dealt with measurement closures rather than positional accuracies.

4.7 Standards for Database Documentation

As described in the Landnet Final Report, corporate data consists of any categories of land-related data utilized (or which could be utilized) by two or more public sector agencies. Regardless of which databases are finally judged to be corporate from a Provincial perspective, there is an existing and growing need for a common standard of database documentation. As pointed out earlier, the quality and availability of current database documentation among departments ranges from extremely detailed and complex (as in the case of PATS) to simply a listing of data fields in the database. Particularly in cases where the database is created and maintained by a very small group of specialists, the Project Team often found little or no supporting documentation covering overall areal extent, data collection or interpretation specifications, updating cycles, attribute consistency, data organization or overall consistency of the information stored in the database.

A common standard for database documentation is not as important in an environment where each specialist group works with their own information exclusively — general descriptions can be provided to new staff members and/or to outside users as the need arises. However, in an environment where the Provincial Government, as a whole, wishes to reduce the amount of redundant data collection, verification and maintenance and wants to increase the integration potential of different databases, then consistent database documentation is an important early step.

COMPONENTS:

- Summary documentation giving the basic characteristics of the data, extent of coverage, intended purposes, currency & updating cycles, and contact persons.
- Detailed documentation concerning:
  - Lineage of the data.
  - Positional accuracy.
  - Attribute reliability.
  - Logical consistency.
  - Completeness of the database.

BENEFITS:

- Would ensure that at least a minimum level of documentation was completed for all databases.
- Would provide a single source of information for users trying to identify and qualify potential sources of information.
- Would reduce the amount of confusion and speculation over existing databases.
- Could limit the amount of potential misinterpretation or misuse of information stored in databases that are unfamiliar to an outside user.

CANDIDATE STANDARDS:

- No standards formally developed.
- Examples of types of documentation are included in Appendix III.

4.8 Standards for Interchange of Land and Water Information Between Different GIS

Communication standards exist on a number of different levels and a detailed discussion of these is beyond the scope of this report. Graphics interchange standards — which allow the transfer of graphics and feature code information from one brand of GIS to another — were discussed in Section 3.0 and in greater detail in Appendix II. It is important to heed international developments in this area and to avoid the temptation to independently develop a local interchange standard.

COMPONENTS:

- See Section 3.0 and Appendix II.

BENEFITS:

- Reduces the number of inter-system or interchange-format translators required by government agencies, contractors and the general public.
- Minimizes the systems support effort involved by each of these groups when systems and/or formats are upgraded.

CANDIDATE STANDARDS:

- See Section 3.0 and Appendix II.

5. STANDARDS IMPLEMENTATION CONSIDERATIONS

5.1 A "Standards Model" for Land & Water Information

Given current trends toward more end-user computing in line departments, there is little chance that all provincial land-related information will ever reside in a single database. While requests for data may be made through a single source, sub-queries will be directed to a number of different databases, and the resulting responses re-accumulated and passed back to the user for evaluation and action. In its envisioned form, the user would request information related to digital maps, records or images based on his/her own "view of the world" in terms of feature definition, representation, etc. Similarly, output from the network would be defined and represented in accordance with the user's own specifications and practices. Ideally — whether the required information is gathered from the database at hand or retrieved from a number of different outside systems — *the whole process should be transparent to the end-user*. While technology will likely provide the tools to accomplish this, agreement on a minimum level of common standards is essential in order to compile, correlate and make meaningful on-line use of the information contained in different information systems.

While standards are desirable in order to exchange data between various groups of users, they need not be imposed in all situations. Users do not necessarily need to standardize on one system, format, or set of specifications across a jurisdiction. Rather, emphasis should be placed on establishing workable standards which represent a common "middle ground" to all departments. For example, as illustrated in Figure 4, organizations *B* and *C* already have well-developed systems and standards in place and are already compatible with one another. The implementation philosophy suggests that these organizations should be able to continue to work with each other using these standards. However, when dealing with other departments or with the general public, they should express themselves using the provincial standards. Organization *A* , on the other hand, may simply subscribe to the Provincial standards from its inception. Note, however, that any contact with the general public — either through accepting input from contractors or for distributing information to outside customers— should make use of Provincial standards.

For example, because they share a common land-cover classification system and compatible GIS software (Arc/Info), forestry companies exchange graphics and text files with DNR as required without going through any intermediate format. On the other hand, if DNR wished to obtain map files from LRIS, then an interchange format would have to be employed to pass graphics files and attribute information between the two systems. Since the classification systems are not alike, documentation describing the classification breakdown would also have to be prepared by LRIS and assessed by DNR before any efficient transfer of data into the target system could take place.

In the interests of minimizing the resources required in this matter, the number of 'common standards' recognized by the Province should be kept to a minimum. While interchange formats largely take care of major concerns over system incompatibility (within the limitations discussed in Section 4.8), they do not address differences in feature/entity definition, data classification, spatial referencing or cartographic representation. In order to achieve the level of transparency discussed earlier, common standards are also required in these areas.

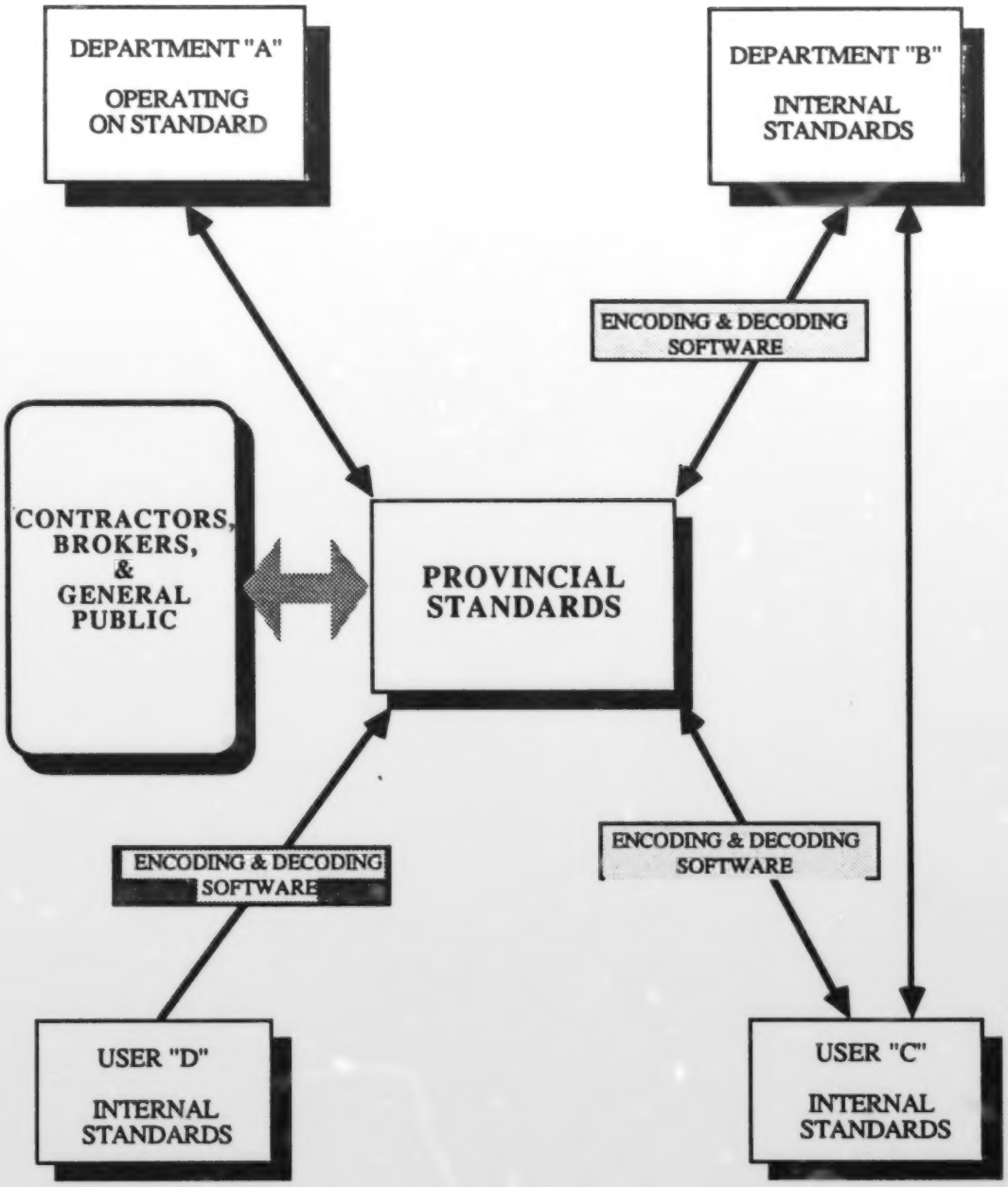

FIGURE 4.

COMMUNICATION IN A STANDARDS ENVIRONMENT
(Extended from Canadian Council on Surveying and Mapping, 1984)

5.2 Implementation Philosophy

Standards — like comprehensive base mapping — represent basic information infrastructure and, as such, are a tough concept to sell. The front-end costs in experienced manpower and sophisticated equipment are high, while the benefits (however substantial) accrue at a later date and are difficult to isolate using current methods of departmental cost-accounting.

The Province has a wide spectrum of options in addressing standards implementation at the executive level. At one end, it could take no direct action and simply allow line departments to continue collecting information independently. At the other, it could formulate comprehensive standards for database content, classification, attribute reliability, spatial consistency, timeliness of information, allowable updating cycles, etc. and compel all departments to collect and maintain their information in accordance with these standards.

Adopting either extreme will produce disappointing results. The 'laissez-faire' approach would produce no improvement in existing practices and would do nothing to promote data sharing. Experience in the United States has shown that the other extreme is equally frustrating — the interdepartmental negotiations required can be protracted, and the cost to modify existing records to meet new standards may be prohibitive. In short, neither extreme will produce the desired effect. Developing appropriate standards involves balancing the goals of increased flexibility and wider use of corporate information against the prospects of tighter centralized control and higher overhead costs. *The Provincial Government must find a middle ground from which to base its standards development efforts.*

To date, the only proven incentive to ensuring a standard is adopted has been to have one in place. In order for the standards to be adopted in practice, the standard must possess four attributes:

- RELEVANCE: There has to be a genuine need (immediate or impending) for the standard.
- EXISTENCE: Database standards cannot be employed if existing data does not cover the area of interest. If a particular base mapping series or property database is to be designated as a provincial standard, then full geographical coverage should *at least* be available in the areas of greatest use. In addition, the Province should commit itself to complete Provincial coverage as quickly as possible.
- SUPPORT: Standards should be easy to access and use. Detailed, understandable documentation explaining the contents and characteristics of the database and/or specification should be prepared and kept up-to-date. If possible, meaningful default models and examples should accompany the documentation. To avoid confusion, there should be one definitive source of technical support for the standards in the Provincial government.
- FLEXIBILITY: While common standards are desirable for the exchange of data between different groups of users, they need not be imposed in all situations. An agency should not necessarily have to use Provincial standards when exchanging information with another organization sharing the same internal standards for data classification and operating on a similar system. *However,* the agency must be able

to relate or transform its information to the common standards when exchanging digital data with an unlike system or when distributing information to the general public.

Rather than developing standards from first principles, every effort must be made to make use of standards implementation and development models from other jurisdictions. Efforts are now underway in the federal government and the provinces of Alberta, British Columbia and Ontario — as well as in the United States and Australia — to formalize standards. More importantly, standards development and implementation must take into account the GIS and information management expertise in user departments which have already implemented systems of their own (including the Department of Natural Resources, the Department of Agriculture, and the LRIS operations in Amherst and Summerside).

5.3 Proposed Framework

Figure 5 illustrates specific GIS standards and implementation issues to be addressed by the Provincial Government. The standards for spatial referencing, parcel definition, GIS data interchange, data classification and object representation are presented in order of priority against the backdrop of common standards for relevant terminology and database documentation. The respective characteristics, benefits and current alternatives for each of these standards have been discussed in Section 4.

Many of the activities shown in Figure 5 are components in programs already underway in New Brunswick. In an environment where central government wishes to assume a leadership role, the overall Provincial interest must be balanced against the program requirements of these individual departments.

Even with cooperative development and testing of Provincial standards, some overlap is bound to remain — existing departmental and/or professional standards have been developed over a long period and influence the activities of many other groups both inside and outside the Provincial Government. In the short term at least, the existence of these localized standards must be a factor in the development of their Province-wide counterparts. To maintain balance and momentum a clear distinction must be made between information of a "corporate" nature and data unique to a particular department.

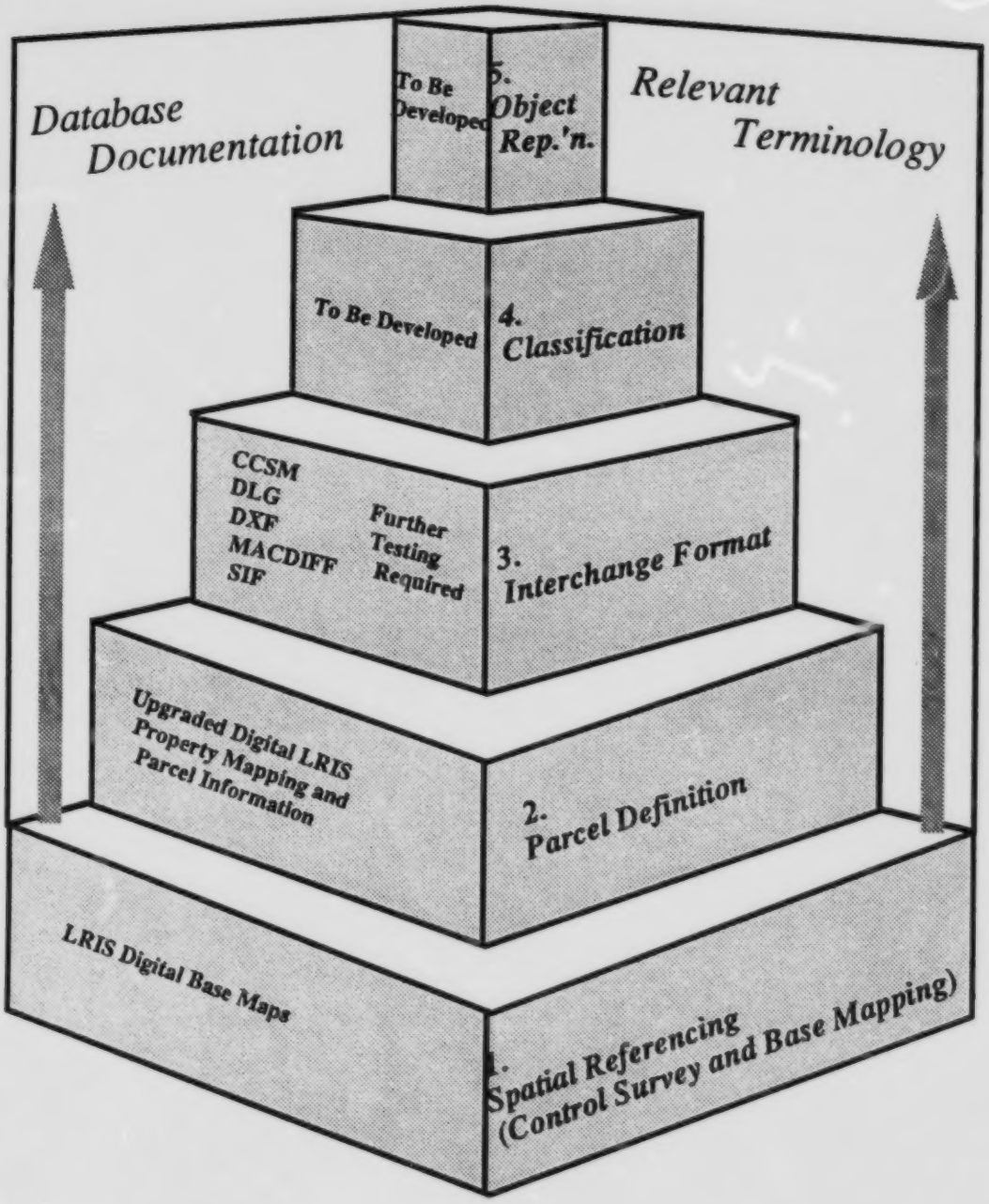

FIGURE 5

BUILDING BLOCKS FOR STANDARDS IMPLEMENTATION

5.4 Recommendations

These recommendations are made as a result of the investigations undertaken during the course of this project. They are based on the working principles put forth earlier and are meant to be implemented in the order listed.

5.4.1 *ESTABLISH AN ORGANIZATIONAL FRAMEWORK*

• *Designate a lead agency to coordinate and facilitate standards development, testing, implementation and maintenance.* Since the preliminary version of this Report was submitted, the Geographic Information Corporation (GIC) was announced and has now been given responsibility for development and implementation of selected land- and water-information standards in New Brunswick.

• *Establish an inter-departmental advisory committee under the direction of the lead agency to guide the implementation of the recommendations detailed in this report.* This group could be a sub-committee of GIC's Advisory Board or it could stand on its own. In either case, it should be established to provide both technical advice and departmental input during the development and/or adoption of standards. As implementation is completed, Committee members could provide input on the on-going adjustments or revisions required. As required, short-term working groups could be formed by this Committee to provide detailed technical advice on implementing the standards outlined in the following sections.

• *Determine an incentive scheme which encourages the adoption and use of Provincial standards.* In the United States Federal Government, efforts have been underway for more than 6 years to develop the standards required to exchange data between different federal departments. Although the government has made a significant investment in time and personnel, these efforts have been stalled due largely to jurisdictional and mandate disputes between different departments. In 1988, a new committee suggested that the federal Office of Management and Budget implicitly endorse the standards by linking the program funding that "feeds" database development and maintenance to the acceptance of selected minimum standards.

Policy decisions and/or regulations notwithstanding, the best motivation for adopting a standard is still *its ability to do the job required.* (Refer to the four prerequisites discussed in Section 5.2.)

5.4.2 *ESTABLISH INTERIM GUIDELINES DURING THE STANDARDS DEVELOPMENT PERIOD*

While government-wide standards are under development, *specify that documentation of de facto standards be improved and maintained by the agency most closely connected to the database in question.* If cooperative updating is agreed upon as a goal, it should be specified that any return of updated data or transfers of the same data to a third party will be carried out in the same format as the original file. This concept was introduced successfully as an intermediate standards measure for the transfer of parcel attribute data in the Queensland, Australia land information network.

Constraints or Concerns:

- The agency that may appear to be the most appropriate custodian is not always in a position to assume that role. For example, it may not be in a position to provide Province-wide coverage or may not be in the best position to keep the information up-to-date.
- Logical candidates for custodianship may not have the mandate to collect data appropriate for use by other disciplines. For example, agricultural soils data may not be entirely suitable for forestry-related analysis, environmental assessment, or highway engineering.
- Organizations are motivated to make use of outside data because it is cheaper than collecting it themselves. If updates or corrections must be recorded to the original standards, then it may cost the department more to use this information than to collect it from scratch. Unless there is some way of indicating the net benefit to the Province, current departmental accounting practices would not recognize the "savings" accrued through using information from outside agencies.

5.4.3 *DEVELOP STANDARD TERMINOLOGY*

Begin immediate development of a standard set of relevant terminology. This is necessary to make sure that subsequent planning and implementation is carried out in an environment where all contributors share a common understanding of important terms. Rather than starting "from scratch", material from glossaries of GIS- and LIS-related terms already in place in Alberta, the Federal Government and the United States can be modified to meet Provincial requirements. Definitions should be refined and localized where necessary, in order to produce the final document.

5.4.4 *DEFINE A STANDARD FOR SPATIAL REFERENCING AND BASE MAPPING*

Confirm a single set of standards for spatial referencing and limit the number of reference systems now in use. The Provincial Land Information Policy submitted to Cabinet has proposed the LRIS Base Map as the standard for base mapping in New Brunswick. Adoption of this base also implies a single Provincial standard for geo-referencing (x,y coordinates based on the N.B. Stereographic map projection and tied to the ATS'77 reference ellipsoid). While not explicitly indicated, the LRIS conventions for symbology at resource and urban scales could be defined as the interim provincial standard for map symbology. This would provide beginning users (and the general public) with a general set of already-recognizable symbols and (potentially) minimize the amount of time new users would have to spend creating their own symbol libraries. In addition, it would act as a "first-approximation" of the final standard and would give more experienced users the chance to identify possible conflicts within their own map specifications.

Constraints or Concerns:

- The Province is not completely covered by LRIS digital base mapping at this time, and this lack of complete coverage could potentially result in the use of other digital databases.
- A number of agencies use EMR 1:50,000, 1:250,000 and 1: 1 million-scale conventional base mapping because they either: (a) are accustomed to the smaller-

scale and the convenience of representing a large area on a single map sheet; (b) interact with federal agencies requiring geo-referenced data in UTM coordinates related to the NAD'27 datum; or (c) have built up an investment of thematic data related to the EMR base and believe it would not be cost-effective to break with traditional practices. Such agencies may be extremely reluctant to change their practices.

- Mathematical transformations between the EMR and LRIS bases are possible and the UTM coordinates of single points can be transformed to their corresponding stereographic version using software available through LRIS. However, organizations should be aware that: (a) this process becomes expensive when a great number of point-transformations are required; and (b) the corresponding map locations may not appear to match precisely due to differences in the respective interpretation, dates of photography and levels of content and symbology.
- According to experiences in the N.B. Department of Agriculture, the LRIS map symbol libraries for resource mapping are built on an earlier PDP version of CARIS which is incompatible with the VAX- and UNIX-based versions now being sold. This means that further work may need to be carried out by either USL or LRIS before the standard can be implemented in practice on other provincial CARIS systems. Further investigations are required to determine: (a) whether it is a general problem or only a problem on the LRIS Carto files; (b) the full nature and extent of the problem; and (c) the costs of completing the symbol libraries for other CARIS users.

5.4.5 *AGREE ON A STANDARD DEFINITION AND IDENTIFIER FOR PARCELS*

Confirm a standard for "parcel" definition and boundary representation. Policy proposals to Cabinet have also recommended the LRIS Parcel Identifier (PID) number as the standard parcel number to be used in all Provincial property-related information systems. This implicitly recognizes the LRIS definition of a parcel as being the Provincial standard as well. In the past, differences in parcel definition between LRIS, the PPDC and the Assessment Branch has caused problems in the consistent collection and sharing of parcel information. However, since: (a) the LRIS Property Mapping and the corresponding LRIS Parcel Index now cover the entire Province; (b) verification and updating mechanisms between LRIS and the PPDC have improved; and (c) LRIS and the Assessment Branch have agreed to co-operate on a data integration pilot project in Saint John, there should be little institutional resistance to the implementation of this standard.

Constraints or Concerns:

- Property lines depicted on LRIS maps are not guaranteed to be accurate or complete in all cases. Information now possessed by government agencies (e.g., the PPDC) and large private landowners (e.g, J.D. Irving and Fraser Co. Inc.) will have to be employed to improve the quality and representation of information contained in the LRIS products. While some *ad hoc* arrangements are now in place, there are no consistent policies or mechanisms governing the cooperative upgrading and updating of property mapping.

- The LRIS definition of a parcel usually— but not always — matches the legal definition for the property. Exceptions include leased properties, Crown land, and expropriated "sliver-parcels" adjacent to highways.
- Parcels contained in the Assessment Database and the PPDC, among others, do not always have a 1:1 correspondence with the LRIS parcel. If the intent is to be able to access information in these databases using the LRIS PID as a primary access key then, the respective relationships between the two sets of parcels will have to be determined and accounted for in each case. Such efforts are now underway between LRIS and the Assessment Branch.

5.4.6 *ARRANGE FOR FURTHER INVESTIGATIONS BEFORE SELECTING A STANDARD INTERCHANGE FORMAT FOR GIS DATA TRANSFER*

After completing a detailed technical evaluation of existing formats, adopt a single standard interchange format for data exchange between unlike systems; contract submissions; and data distribution.

Operationally, it makes sense for users of similar systems to pass files to one another in original system format. This is a fairly common practice in the Provincial forestry sector. Since file transfer between unlike systems is still a relatively uncommon, although increasing practice, the Province still has the opportunity to limit the number of interchange formats in use without a great deal of protest from user departments. Provincial land information policy proposals have recommended that all subsequent GIS purchases by Provincial Government departments be limited to CARIS software. However, since there are already systems from at least five different vendors now in place in New Brunswick, a common interchange format is still desirable.

It is recommended that the relative performance and flexibility of the various interchange formats be examined before standardizing on one. Since major GIS vendors are committed to supporting most of the major interchange formats discussed in Section 3.0, the acceptance criteria should be based on relative performance (in both batch and on-line mode), flexibility, and potential support by lower-end systems which may be used by the general public. It is important that this project be undertaken, and a single interchange format adopted, before user departments specify (or develop) their own interchange formats for contract submissions and data distribution. This testing could be carried out by either internal staff or through contract. In either case, the investigation should employ the expertise and facilities already in place at CanLab-INSPIRE at UNB.

5.4.7 *DEVELOP A STANDARD FORMAT FOR DATABASE DOCUMENTATION.*

Overview documentation should be contained in 2-3 pages and include point-form information concerning database description, data sources and estimated reliability. Two good examples of the points to be covered — one from New Zealand and one from a 1984 federal government database inventory — have been included in Appendix III.

Detailed documentation will provide interested users with enough information to evaluate the fitness of the data for a particular use. The proposed U.S. standard for digital cartographic data presented extensive guidelines for detailed documentation, and this standard has since been adopted by both the State of Wisconsin and Statistics Canada. To

implement this portion of the standard, producers must include rigorous and quantitative information available on all database components.

5.4.8 *PREPARE DATABASE DESCRIPTIONS*

Require all departments to prepare database documentation based on the Provincial standard adopted above.

By this point, the foundation sub-system (see Figure 2) has been standardized and a communication standard has been adopted to facilitate communication between theme databases. However, unless a set of capsule descriptions of the databases maintained in different departments is prepared, compiled and maintained, most users will continue to be unaware of the types of data already available. Unless detailed documentation is available, potential users will be unable to make informed decisions about the dataset's potential fitness for use in their application. By documenting these databases in a consistent manner, users will have a greater awareness of the information available and will be able to judge the potential applicability of databases within their own terms of reference. If suitable, they can then physically relate the data to their own classification scheme and decide whether it is appropriate to their needs. This is still far from the transparent on-line network desired since differing standards for classification and representation will still exist, but significant improvements to the current situation will have taken place.

Constraints or Concerns:

- The fact that so little documentation currently exists for Provincial databases indicates that departmental staff may not have the time to create or maintain this documentation in the manner necessary. Incentives will have to be established to ensure the documentation is completed to the standard, verified and kept up-to-date.
- An example of completed detailed documentation is included in Appendix III.

5.4.9 *DEVELOP A STANDARD FRAMEWORK FOR DATA CLASSIFICATION*

No single classification breakdown emerged as a "standard" from the site interview process. Therefore the development and adoption of a standard will depend upon whether or not a common framework can be developed which can accommodate and categorize the types of data maintained in various databases. The CCSM framework for feature classification is a good example of an open-ended standard, although it would have to be extended to accommodate parcel ownership and land use information.

Once the framework was developed, organizations with information systems already in place will need to determine the level of effort required to relate their own information to this framework. While their internal methods of database organization will ultimately depend on departmental data retrieval and reporting requirements, it will be necessary for them to be able to relate their data to this standard framework for interchange purposes.

5.4.10 *STANDARDIZE OBJECT REPRESENTATION*

Adopt the LRIS specifications as the standard for object representation until more user feedback can be obtained.

A robust standard for object representation is difficult to implement without having a clear picture of potential GIS applications in their respective organizations. Working from the classification framework developed above, a defined number of alternative representations must be developed for specific feature classifications within a given scale range. Initially, the LRIS specifications can be employed as the interim standard until there is more feedback on the strengths and limitations of that particular series with respect to most applications. Once the feedback is obtained, the standard can be revised as required.

5.5 Balancing the Needs of Experienced and Non-Experienced Users

Based on the results of the site interviews, prospective GIS users are receptive to a knowledgeable third-party providing guidance and examples dealing with standards implementation. More experienced users, on the other hand, have already made a substantial investment in developing their own standards and procedures. While these users usually endorsed the need for common standards, they were more guarded about full support until the "standards" in question had been identified. *To ensure smooth implementation, the requirements of prospective users must be balanced against the existing financial and intellectual commitments of organizations more experienced with GIS.*

To address the needs of experienced users:

- Make sure that senior technical and operations personnel from organizations with the most experience are fully involved in selecting and/or planning new standards.
- Agree on a "minimum level" of standards conformity to be adopted by all departments and set tentative milestone toward reaching this goal.
- Agree on circumstances where internal specifications (rather than Provincial standards) can still be employed when dealing with third parties.
- Work with management and operations personnel in each organization to determine the respective ramifications of implementing the selected standards.
- Develop cost and scheduling budgets to estimate the resources required to revise their existing databases and procedures to accommodate these standards.
- After reviewing cost and scheduling estimates from all affected organizations, allocate the resources as funding permits and adjust schedules where necessary.

To address the needs of prospective users:

- Ensure that documentation of current standards is complete and available.
- Prepare documentation describing how these standards can be most effectively employed.
- As standards for classification and object representation are developed, models of coding values and representation should be prepared and tested wherever possible. The resulting examples should be considered as "default values" suitable as a starting point for more focused development in the user's applications.

- Prior to hardware or software acquisition, demonstrate how standards can be applied to program requirements. This may be accomplished through special user-group meetings or individual sessions with the agency in question.
- Ensure that any land-related database or GIS pilot projects in government departments ascribe to the adopted standards.

SELECTED REFERENCES

CCSM (Canadian Council on Surveying and Mapping). 1984. Data classification, Quality Evaluation and EDP File Format. National Standards for the Exchange of Digital Topographic Data, Technical Committee Volume 1. Ottawa.

CCSM (Canadian Council on Surveying and Mapping). 1984. Topographic Codes and Dictionary of Topographic Features. National Standards for the Exchange of Digital Topographic Data, Technical Committee Volume 2. Ottawa.

CCSM (Canadian Council on Surveying and Mapping). 1984. EDP Standards Applied to Digital Topographic Data. National Standards for the Exchange of Digital Topographic Data, Technical Committee Volume 3. Ottawa.

Mollering, H., ed. 1984. Digital Cartographic Data Standards: Examining the Alternatives, Issues in Digital Cartographic Data Standards. Report No. 4, Columbus Ohio: National Committee for Digital Cartographic Data Standards.

Mollering, H., ed. 1984. Digital Cartographic Data Standards: An Interim Proposed Standard, Issues in Digital Cartographic Data Standards. Report No. 6, Columbus Ohio: National Committee for Digital Cartographic Data Standards.

Mollering, H., ed. 1984. Digital Cartographic Data Standards: A Report on Evaluation and Empirical Testing, Issues in Digital Cartographic Data Standards. Report No. 7, Columbus Ohio: National Committee for Digital Cartographic Data Standards.

(Additional References in Appendix II)

APPENDIX I

LIST OF AGENCIES INTERVIEWED

APPENDIX I

LIST OF AGENCIES INTERVIEWED

1. PROVINCIAL GOVERNMENT

DEPARTMENT	DIVISION
(i) Agriculture	On-Farm Soil Survey
	Farm Land Identification Program
	Livestock Feed Initiative
(ii) Justice	County Registry Offices
	Albert Co. Land Titles Office
(iii) Municipal Affairs & Environment	Assessment (PATS)
	Community Planning
	Water Resources
	Emergency Measures Organization
(iv) Natural Resources	Crown Lands
	Forest Extension Services
	Mineral Resources
	Timber Management
(v) Supply & Services	Buildings and Properties
	Information Systems
(vi) Tourism, Recreation & Heritage	Field Services
(vii) Transportation	Design Branch
	Right-of Way Branch
	Planning Branch

(viii) LRIS	Property Mapping Division Surveys and Mapping Branch Amherst Group

2. FEDERAL GOVERNMENT AGENCIES

DEPARTMENT	DIVISION
(i) Agriculture Canada	Land Resource Research Centre
(ii) Emergency Preparedness Canada	Emergency Preparedness Canada
(iii) Energy Mines and Resources	Geological Survey of Canada Canada Centre for Mapping
(iv) Environment Canada	Inland Waters Directorate CGIS National Parks Directorate Canadian Forestry Service (Canadian Forest Inventory) Canadian Forestry Service (Forest Insect and Disease Survey)
(v) Fisheries and Oceans Canada	Canadian Hydrographic Service Habitat & Resource Development
(vi) Statistics Canada	Geocartographics Group
(vii) Transport Canada	Architectural and Engineering Services

3. MUNICIPALITIES

(i) City of Moncton	Engineering and Works Department
(ii) City of Fredericton	All Departments

4. INDUSTRY

(i) J.D. Irving (Woodlands) Ltd.

(ii) Fraser Company Inc.

(iii) N.B. International Paper

(iv) The NB Telephone Company Ltd.

(v) NB Power

(vi) Real Estate Board of Fredericton Inc.

APPENDIX II

DETAILED DESCRIPTIONS OF INTERCHANGE FORMATS

APPENDIX II

DETAILED DESCRIPTIONS OF INTERCHANGE FORMATS

A. Canadian Council on Surveying and Mapping (CCSM) Standard

1. Introduction

The Canadian Council on Surveying and Mapping (CCSM) Standard was designed as a Canadian national standard for the exchange of digital topographic data. It is the work of a Technical Committee formed from representatives of federal and provincial government departments and the private sector. The Department of Energy, Mines and Resources (EMR) plans to accept maps from contractors only in the CCSM standard starting with the fiscal year 1990/1991. During the year prior to that, EMR will accept maps in either the CCSM or SIF formats. It is not known if other Federal departments will be imposing similar restrictions.

Spatial objects are called **entities**. They are of three types: point, line and area. Each entity consists of a 15-digit identification number, a 12-character feature code, a spatial description and a set of attribute values. A **point** is an entity defined by a triplet of (x,y,z) coordinates. A **node** is a point which is a 0-cell. A **line** is either defined by a string of connected coordinate triplets, or by the parameters of an arc of ellipse or an arc of circle. A **co-located line** is one that shares the coordinates of another line. A **boundary line** is a 1-cell. An **area** is a 2-cell. The terms nodes and boundary lines are not consistently used in the CCSM standard document referenced. A topological node is sometimes implied by a point connected to lines and a boundary line is implied by a line with left and right areas identified.

Attribute values can be imagined to be organized in a tabular form. The CCSM format assumes that the attribute tables have at most 143 columns (**attribute fields**). Each row hence contains the attributes of one entity. There is an **attribute descriptor** for each of the 143 fields. It includes the attribute name, type and length of the field.

Entities are organized into logical volumes, data sets, data groups and data themes. A **logical volume** is a collection of data sets and can span several physical volumes. A tape, for example, is a **physical volume**. There is a file called the VOLUME DIRECTORY FILE at the beginning of each physical volume. A number (PHYSICAL VOLUME NUMBER IN LOGICAL VOLUME) in this file indicates the sequence of the physical volumes in the logical volume.

A **data set** is a collection of data groups in the same geographic area. For example, data belonging to a map sheet can be stored in a data set.

A **data group** is a collection of data themes that could be topologically connected. It is hence equivalent to a planar graph if the data themes are topologically related. For

example, a land parcel layer consisting of nodes, boundary lines and areas form a data group.

A **data theme** is a collection of entities of the same type and share the same attribute fields. For example, all the boundary corners of land parcels can be represented in a point theme. Similarly, all the boundary lines and all the parcels can be represented in a line theme and an area theme respectively.

Whether a data group is topologically related depends upon the nature of the data themes it contains. For instance, a collection of point, boundary line and area themes described in the previous paragraph forms a topologically related data group. However, a collection of two point themes, such as a group of spot heights and a group of triangulation points, which are not topologically related to each other, can also be formed into a data group.

2. Coordinate System

All coordinates are expressed as (x,y,z) triplets. The z-coordinate contains a 0 if it is not used. The coordinates are relative to a single origin defined in the map header and each can be expressed using

- a 15-digit integer (plus a sign),
- a real number with at most 10 significant digits plus two additional digits for the exponent,
- a degrees-minutes-seconds combination for the geographic coordinate system. In this case, three digits are used for the degree, two for the minute and seven for the second.

3. Projections Defined

The following six projections are defined in the CCSM standard: simple latitude/longitude (a trivial projection), UTM, Mercator, Lambert Conformal, Stereographic and Polyconic.

4. Spatial Objects

The CCSM standard supports the following spatial objects:

- **Point** - A point represented by a coordinate triplet, a point identification number, a feature code, an orientation angle and attribute values.
- **Line** - A connected sequence of points. It is associated with a line identification number, the identification of a co-located line, a feature code and attribute values.

- **Arc** - A special line which is either a planimetric arc of an ellipse or an arc of a circle. An arc is defined by the start and end points; the origin and lengths of the primary and secondary axes; and the orientation angle of the primary axis.

- **Text** - A text string of at most 64 characters and is associated with a coordinate triplet, an identification number, a feature code, an orientation angle, justification, font, text height, text width, the text string and attribute values.

- **Node** - A point topologically related to boundary lines. It is organized like a point but with the following additional information: the number of lines attached to this node, the identification of the lines attached.

- **Boundary Line** - A line topologically related to two nodes and forming the boundary of areas. It is organized like a line but with the following additional information: identification numbers of the start node, the end node, the left area and the right area.

- **Area** - A closed polygon. It is associated with an area identification number, a point inside the area, a feature code, the number of boundary lines forming its perimeter, the identifications of the boundary lines and attribute values.

5. Topology Handling

The topological entities supported are the nodes (0-cells), boundary lines (1-cells) and areas (2-cells). Each node is linked to all boundary lines sharing the same node. These boundary lines, however, are not arranged in any particular order around the node. Each boundary line is linked to the two nodes at both ends and it is also linked to the two areas on both sides of the line. Each area is linked to all the boundary lines forming its perimeter. There is no explicit fields in the data records for enclosure relationship.

6. Coding Method

A detailed description of the coding method is given in CCSM [1988, Appendix A]. The following is a summary of the salient features:

The CCSM exchange file is an ASCII file organized into logical volumes, physical volumes, data sets, data groups, data themes and entities. The meaning of these terms has been explained in an earlier section. The entries in the exchange file are not tagged, but their meanings are implied by their position in the file.

There is no header record for a logical volume. However, the VOLUME DIRECTORY FILE preceding each physical volume (e.g. at the beginning of each tape) contains a LOGICAL VOLUME IDENTIFIER and a sequential number showing its position in the logical volume. Other information contained in the VOLUME DIRECTORY FILE include its creation date and the generating agency.

Following the VOLUME DIRECTORY FILE are the data sets. Each data set starts with a DATA SET HEADER RECORD containing the data set name, its creation date, the

number of data groups in this set, parameters related to the coordinate system and information about resolution and accuracy of data.

No header record is used for a data group. Instead, the DATA THEME HEADER RECORD at the beginning of each data theme contains a count of the number of point, line and area data themes in the group. All the data themes belonging to a data group are stored next to each other. Other information contained in the DATA THEME HEADER RECORD include the data theme name, the entity type (point, line, or area), the feature class (a high level code for the features) and the length of entity records in this data theme.

After the DATA THEME HEADER RECORD are the ATTRIBUTE DESCRIPTOR RECORDS for the theme. There are at most 143 of them and each contains a description of the attribute name and the attribute data type to indicate whether it is an integer, a real number, an so on. The attribute values of the entities will be entered, later in the entities records, in the same order as their attribute descriptors.

The entities belonging to a data theme are stored together after their DATA THEME HEADER RECORD. Each entity is composed of a fixed-length record and an optional variable-length record. The length of the fixed-length record is contained in the DATA THEME RECORD. The contents of the entity records has been described in the section on Spatial Objects.

7. Strengths and Weaknesses

The CCSM standard is designed for the exchange of topographic data. It is a simple format but contains sufficient information to define projection parameters, nature and resolution of the coordinate system, position, identification and attribute values of entities. It can also contain topological relationship among the entities.

A minimum number of entities are defined and similar entities (such as points and nodes) share the same data structure. Most of the fields and the records are of fixed length. These help to simplify the reading and writing of the exchange file and improve the readability of the file on a hardcopy. The concept of a co-located line is useful in a situation where the same geometric line is shared by several different linear features.

The fixed-length fields and records sometimes require unnecessary storage. For example, coordinates are always expressed as triplets, even for entities that does not require z-coordinates. This makes the coding of contours rather awkward: the height of the contour is stored as an attribute value and every z-coordinate of the points on the contour line are forced to be zero.

Other weaknesses include the lack of enclosure relationships, the inability to either form compound objects,or bind a text string along a curve and the lack of facilities to handle symbology. However, these problems can always be resolved through the use of the attribute records, although this practice would violate the spirit of standardization. It should be noted that the CCSM standard is still evolving and new spatial objects could be included from time to time. The coding method seems flexibility enough to accommodate future enhancements.

8. Reference

CCSM (1988). *National Standards for the Exchange of Digital Topographic Data, Volume I/Part III, Standard EDP File Exchange Format for Digital Topographic Data*, Version #1.0, Canada Centre for Mapping, Energy, Mines and Resources, Ottawa, Ontario, Canada, May.

B. Digital Line Graph (DLG) Format

1. Introduction

The Digital Line Graph (DLG) format was developed by the United States Geological Survey (USGS) as a digital cartographic data standard. This standard, originally designed for the distribution of USGS digital maps, is now widely used for the exchange of data between different mapping systems.

There are three levels of the DLG format. The **level one format (DLG-1)** is designed for the quick transfer of raw, unstructured data to be further processed by the target system, or for the production of cartographic products of low to medium quality. The **level two format (DLG-2)** supplements the level one format with extensive attribute coding and is suitable for the production of high quality cartographic products. The **level three format (DLG-3)** in addition supports topology. It is not expected that data collected with level one and level two specifications can be upgraded easily to meet the level three standard. However, level one and level two formats can be derived from the level three format. The DLG-3 format will be described in this report because it is the most complete of the three.

A DLG file is an ASCII text file and has two formats of distribution: **standard** and **optional**. The optional format supports coordinates in ground units and provides more explicit topological relationships. Being more compact, the standard format is suitable for data archive whereas the optional format is suitable for data interchange.

The basic spatial objects in a DLG file are two-dimensional and are called **data elements**. Each element is assigned an **internal identification number**, an optional **text string** and can be assigned none or many **attribute codes** (or **feature codes**). A DLG attribute code is hierarchically structured into two fields: a three-digit **major code field** and a four-digit **minor code field**. An attribute code for a data element is not necessary if the nature of that element can be derived from other information. For example, a line that is a boundary between parcels can be derived from the attribute codes of the two neighbouring parcels.

At the DLG-3 level, the topological elements are the **nodes** (0-cells), the **lines** (1-cells) and **areas** (2-cells). The area outside of the map border has an attribute code of 000 0000 and is the first area in a DLG file. The nodes and lines are associated with

coordinates but areas are not. Using the standard DLG distribution format, only lines contain explicit topological linkages which are: the start and end nodes as well as the left and right areas. Other topological relationships, such as the lines forming an area, must be derived indirectly. There is currently no provisions in the DLG format to handle island information in the standard format. Using the optional format, all lines sharing the same node, all lines forming the perimeter of an area and island information are recorded.

In the case of a network in which enclosed areas are not of interest, only two areas are identified: the outside area and the inside area called the **background area**. All lines on the network, except those along the border of the map, are considered to be enclosed by the background area and hence have the same area identification number on both sides of the line.

The nodes, lines and areas that are topologically related to each other form a **data category**. A DLG file can contain several data categories.

A DLG file contains a number of **header records** describing the important parameters of the map, including: name, date and scale of map, projection parameters, ground unit, resolution, transformation parameters from internal file unit to ground and registration points.

2. Topology Handling

At the DLG-3 level, all data elements are topological elements. They are the nodes (0-cells), the lines (1-cells) and the areas (2-cells). Several planar graphs (called categories) can be stored in a single DLG file. Using the standard DLG distribution format, topological relationships are explicitly coded for the line and includes:

- the nodes at the two ends of the line, and
- the areas on both sides of the line.

The optional DLG distribution format provides the following additional explicit relationships:

- the lines sharing the same node,
- the lines forming the perimeter of an area, and
- islands within an area.

3. Coordinate System

There are two formats for the encoding of points which are two-dimensional in a DLG file. The standard format is based on a Cartesian system with the origin either at the lower left corner or at the centre of the map. Each coordinate is expressed as a signed 5-digit integer ranging from -32767 to +32767 and the unit is 0.001 inch at the map scale.

The optional format represents each coordinate as a real number with 12 significant digits, two of which are after the decimal point. These coordinates are in ground units, which is usually metres in the UTM projection.

4. Projection Defined

The UTM projection is defined in the DLG format.

5. Spatial Objects

The DLG standard supports the following spatial objects:

- **Node** - The endpoints of a line. Associated with a node record are: an internal identification number, the coordinates of the point, attribute values and a character string.
- **Line** - An ordered set of points. Associated with a line record are: an internal identification number, the coordinates of the points on the line, the starting node, the ending node, the left area, the right area, attribute values and a character string. A line contains from 2 to 1500 points.
- **Degenerate Line** - A line that starts and ends at the same node, has zero length and is not attached to any other line. It can used to represent a point feature.
- **Area** - A closed polygon. Associated with an area record are: an internal identification number, the coordinates of an arbitrary point not necessarily inside the area, attribute values and a character string. Using the optional DLG distribution format, the number of islands is also included.

6. Coding Method

A DLG file is comprised of 8-bit ASCII characters. The logical records are fixed length and are 144 bytes long in a standard file but 80 bytes long in an optional file. In an optional file, only the first 72 bytes are used for data whereas the other 8 bytes contain a record sequence number.

Several record types are recognized. For instance, there are record types for line description records, coordinates records and attribute code records. The data belonging to one record type could span several logical records and fillers are often used within a record type for padding. Data fields within a record are not tagged, but the meaning of a particular item is implied by its position within a record. The record types do not carry identification codes except the node, line and area types which are labelled as "N", "L" and "A" in the element-type fields. In general, the records are ordered in the following sequence in the standard DLG distribution file:

Header records
- Records containing DLG identification information
- Records containing projection information and registration points
- Records containing data category identification

Data records (all data records of a data category are grouped together)

- Node records (all node records within a data category are grouped together)
  - Node description
  - Attribute codes
  - Text string
- Area records (all area records within a data category are grouped together)
  - Area description
  - Attribute codes
  - Text string
- Line records (all line records within a data category are grouped together)
  - Line description
  - Coordinates
  - Attribute codes
  - Text string

The records in an optional DLG distribution file are similar to that of a standard file and are ordered in the following sequence:

Header records
- Records containing DLG identification information
- Records containing projection information
- Records containing control points
- Records containing data category identification

Data records (all data records of a data category are grouped together)

- Node records (all node records within a data category are grouped together)
  - Node description
  - Node-to-line topological linkage description
  - Attribute codes
  - Text string
- Area records (all area records within a data category are grouped together)
  - Area description
  - Area-to-line topological linkage description
  - Attribute codes
  - Text string
- Line records (all line records within a data category are grouped together)
  - Line description
  - Coordinates
  - Attribute codes

Text string

7. Strengths and Weaknesses

The DLG standard is designed for the exchange of cartographic data. Enough information is carried in the file to define projection parameters, nature and resolution of the coordinate system, position and identification of the data elements. An interesting feature of this standard is the support of three levels of data structure. There are also two formats for data distribution. The standard format is more compact and is designed for data archiving. The optional format contains more information and is the chief one used for data exchange.

Simplicity is one of the attractions of the DLG format. Only a small number of spatial objects are defined: nodes, lines, areas and degenerate lines (isolated points). The files are based on ASCII codes and on fixed-length logical records.

Another strong point of the DLG format is its rich topological content. The highest level data structure, DLG-3, contains many node-to-line, line-to-polygon and polygon-to-polygon relationships. The number of topological relationships defined is the highest in all the standards examined in this report.

Only planimetric maps can be represented in a DLG file because only two-dimensional points are used. There is no mechanism for forming compound objects. The ability to handle attribute information is also limited. Each data element can be associated with a seven-digit attribute code which is really a feature code.

8. References

van Roessel, J.W. and S.W. Doescher (1986). *Vector Data Structure Conversion at the EROS Data Center*, Open-File Report 87-309, Department of the Interior, Geological Survey, Sioux Falls, South Dakota.

USGS (1983). *USGS Digital Cartographic Data Standards. Digital Line Graphs from 1:24,000-Scale Maps*, U.S. Geological Survey Circular 895-C, U.S. Geological Survey, Virginia.

C. ASCII Drawing Interchange (DXF) Format

1. Introduction

The ASCII Drawing Interchange (DXF) Format is used to exchange data between different implementations of AutoCAD and between AutoCAD and other systems. AutoCAD, a general purpose computer aided drafting (CAD) package that runs under the

PC-DOS, MS-DOS and UNIX operating systems, is a product of Autodesk, Inc. of Sausalito, California. The flexibility of this package allows it to be used for map drafting.

Basic objects in AutoCAD are called the **entities**. A complete list of entities is given in Section 2; some examples are lines, arcs and circles. Each entity can be assigned to a layer, given a colour and a line type. A **layer** and a **line type** can be identified by 31-character names whereas a **colour** is represented by a numeric code ranging from 1 to 255. Optionally, colour and line type can be applied globally to all entities of a certain layer.

Compound objects, called **blocks**, can be formed by grouping entities of different layers, colours and line types. Blocks can be associated with one or more **attribute** records. The blocks are referenced by **block names** and can be manipulated in the same way as entities. As such, blocks can be used to define point symbols which are composed of basic entities such as lines and circles. Another way of defining point symbology in AutoCAD is through the use of the **shape file** which is an independent file containing the strokes and other information about **shapes** (point symbols). Many organizations exporting data to non-AutoCAD systems find it easier to transfer point symbology through block definitions which are contained in the interchange file because it requires special software to read the shape files.

A DXF file is an ASCII text file composed of the following four sections: HEADER, TABLES, BLOCKS and ENTITIES. The **HEADER section** contains 88 pieces of information about the file. Some of this information (such as fillet radius and chamfer distance) are used mainly for mechanical drawings; others (such as the highlight mode for selected objects) are related to the editing session.

The following items are of interest to mapping applications:

- Angle base, direction (clockwise or anti-clockwise), unit and precision.
- Date and time of file creation.
- Date and time of file update.
- Unit for coordinates.
- Default layer , colour and line type.

This entire section can be dropped if not used.

The **TABLES section** contains the following tables: LTYPE, LAYER, STYLE and VIEW. The LTYPE table defines the line types (patterns) used, the LAYER table assigns a colour and line type to all entities of the same layer, the STYLE table helps the loading of the font files and the shape files. and the VIEW table defines the view parameters used for an edit session. This entire section, or any of the unused tables can be dropped if not used.

The **BLOCKS section** contains all the block definitions for the map. This entire section can be dropped if not used.

The **ENTITIES section** contains the definition of the spatial objects to be described later. This section is mandatory.

2. Topology Handling

DXF does not recognize topological entities except a polygon defined by a closed polyline. Islands are not supported.

3. Coordinate System

The coordinates in a DXF file can be expressed in absolute, relative and polar forms. Two-dimensional as well as three-dimensional points can be used. Either 16-bit integers (5 significant digits) or 64-bit floating point numbers can be used for the x-,y- and z-coordinates. The number of significant digits is at least 14 in the floating point number format.

4. Spatial Objects

The following entities are supported by DXF:

- **Line** - A straight line segment (vector) defined by two points.
- **Point** - A single point.
- **Circle** - A circle defined by the centre and the radius.
- **Arc** - A part of a circle defined by the centre, the radius and the start and end angles.
- **Trace** - A line (vector) with width defined by four points on the corners of the quadrilateral.
- **Solid** - A solid filled region represented by quadrilateral or triangular sections.
- **Text** - A text string defined by the insertion point, height, text value, rotation angle, relative X scale factor, obliquing angle (for creating the *italic* version), style, a flag to indicate if the text is mirrored or upside down, justification type and alignment point for justification.
- **Shape** - A pre-defined symbol contained in a shape file. A shape is referenced in the DXF file by the insertion point, size, shape name, rotation angle, relative X scale factor and obliquing angle.
- **Block** - A compound object. Blocks are defined in the BLOCKS section of the DXF file and each block is associated with a block name, a flag to indicate if it is associated with attributes and a block base point. A block can be referenced in the ENTITIES section through the use of an INSERT command which carries the following arguments: block name, insertion point, X scale factor, Y scale factor, Z scale factor, rotation angle, column and row counts and column

and row spacings. An INSERT command can be followed by attributes of the block.

- **Attribute Definition** - A text string describing the characteristics of the attribute values. It is used by the system to display the prompt string when requesting attribute input and to decode the contents of the attribute record. It has all the parameters of a text string except text value and carries the following additional information: default value, prompt string, tag string and a flag to indicate if the attribute values are to be displayed, if they are constant, or if input verification is required.

- **Attribute** - Text used to describe blocks. It has all the parameters of a text string except text value and carries the following additional information: attribute value, attribute tag, length of the attribute field and a flag to indicate if the attribute values are to be displayed, if they are constant, or if input verification is required. Several attributes can be associated with a block.

- **Polyline** - A connected sequence of lines (vectors) and arc segments. It can be open or closed to form a polygon. A polyline is associated with a flag indicating if it is closed or if curve-fit information has been added and with default starting and ending line widths. These widths can be overridden by similar parameters associated with each vertex of the polyline.

- **Vertex** - A point on a polyline. It is defined by its location, starting and ending line widths, a flag to indicate if this vertex is a result of interpolation or if a tangent for curve fit purposes has been included for this vertex, bulge and curve fit tangent. Bulge is used to indicate the presence of a circular arc.

5. Coding Method

A DXF file is an ASCII text file. Each item in the file, including section names, table names and values, are tagged with a numeric **group code** occupying three character fields and is right justified. A complete list of these codes can be found in Appendix C of the AutoCAD Reference Manual. Each code is followed, on the next line, by a **group value** which can be a text string, an integer number, or a floating point number. The group code and its group value form a **group** and a DXF file is a collection of groups.

A DXF file is terminated by an end-of-file group:

```
  0
EOF
```

and each section is enclosed by two groups:

```
  0
SECTION

(Groups within a section)

  0
ENDSEC
```

A section name is identified by a code 2 followed by the name. For example, the ENTITIES section marker looks like this:

```
  2
ENTITIES
```

The sections must follow the order HEADER, TABLES, BLOCKS and ENTITIES. Moreover, the tables within the TABLES section must follow the order LTYPE, LAYER, STYLE and VIEW. Entries in the ENTITIES section, however, can appear in any order. The only mandatory section of a DXF file is the ENTITIES section. The following is a small but complete DXF file containing a line and a point:

```
  0
SECTION
  2
ENTITIES
  0                 <-- group code for an entity
LINE                <-- line entity
  8                 <-- group code for layer
LAYER1              <-- layer one
  10                <-- group code for x of start point
556321.43           <-- x coordinate of start point
  20                <-- group code for y of start point
559210.03           <-- y coordinate of start point
  11                <-- group code for x of end point
556425.95           <-- x coordinate of end point
  21                <-- group code for y of end point
559380.43           <-- y coordinate of end point
  0
POINT
  8
LAYER2
  10
551253.45
  20
559184.92
  0
ENDSEC
  0
EOF
```

6. Strengths and Weaknesses

DXF is designed to exchange data mainly for CAD applications. It has a simple format which is easy to understand and use. Blocks can be used to form compound objects such as a spot height which is a composition of a point and a label. Attribute data and their definitions can be included with the file. The fact that a single height value can be associated with an entity makes the representation of contours quite convenient.

A DXF file tends to be long because of the need for a tag (or group code) for every single value. A hardcopy version of it is easy to read but will occupy many pages since every group code and group value must start on a new line. There is a binary version of the DXF file called the DXB file which is more compact but is not designed for inter-system communication.

The line entity is not very useful for digital maps because each one is a vector. Polylines are more suitable for irregular curves, but the present format requires a bulge value and vertex flags for each point. The two pieces of information are only necessary for lines containing arc segments or for lines to be interpolated.

A text string can be rotated as a whole in DXF, but individual characters within the string cannot be rotated or positioned. This makes the binding of a string along a curve (such as a river) difficult.

The header section does not have room for mapping information such as projection parameters. It does not provide a comment field for informal documentation.

7. Reference

Autodesk, Inc. (1986). *The AutoCAD Drafting Package Reference Manual*, Publication TD106-009, June 16, Autodesk, Inc., Sausalito, California

D. Map and Chart Data Interchange Format (MACDIF)

1. Introduction

The Map and Chart Data Interchange Format (MACDIF) is based on the Map Data Interchange Format (MDIF) developed by the Ontario Ministry of Natural Resources in conjunction with IDON Corporation of Ottawa. Other Canadian Federal Departments, in particular the Canadian Hydrographic Service, have contributed to the design of MACDIF.

In MACDIF, the basic graphics entity is called a **segment** upon which **features** are constructed. In general, the segment record carries the coordinates while the feature record carries other descriptions of the spatial object. A segment is uniquely identified by a

number and contains the geometric definition of features. A segment can be real or virtual. A **real segment** defines the outline of a feature, but a **virtual segment** is an imaginary line used to close an otherwise open feature. For example, the intersection of four roads in a large-scale map can be delineated by a rectangle composed of four virtual segments. A **neat line** is a special segment which forms the boundary of the map. Like other segments, neat lines are either real or virtual.

Segments are grouped to form features. Alternatively, a segment can belong to several features. This is the case when several features (such as a river and a parcel boundary) share the same location. A **feature** is identified with a unique number and is associated with a feature code stored in two parts, a reference point, a bounding rectangle defined by the minimum and maximum coordinates, the segments forming the feature, an **annotation name** and its **alternate**, attribute values and others.

MACDIF provides the mechanism for tagging each feature with an alphanumeric **feature type** (or feature code) stored in two parts: the **feature type** and the **feature sub type**. However, MACDIF does not provide a standard for assigning feature types. Instead, it allows the user to reference an established **feature code table**. The tables recognized are those from the Canadian Council on Surveying and Mapping, the United States Defense Mapping Agency, the American Congress on Surveying and mapping and the Canadian Hydrographic Service.

A feature belongs to one of three classes: point-oriented, linear and areal. A **point-oriented feature** is associated with a segment containing a single point. A **linear feature** is associated with a segment containing multiple points. It could be a line, a circle, a circular arc, or a connected series of lines and arcs. An **areal feature** is associated with multiple segments forming a closed loop. It is also possible for an areal feature to be composed of several closed loops either nested within each other as islands, or physically disjoint to form a non-contiguous feature. An alternate method of forming polygons is provided by the topological definition mechanism of MACDIF. Features can be combined into **data groups** which are identified by numerical codes and associated with text descriptions. The main purpose of a data group is not to form compound objects, but to form a layer of features.

Symbology for the features is carried in a separate section of a MACDIF file. The symbology section contains the name of a reference symbol table, rotation and scale of point symbols and characters, patterns used for line symbolization and area filling and the position at which the annotation and alternate names are to be displayed.

A MACDIF file can contain an **overall map** and several **sub-maps** in the form of overlays or inserts. An **overlaying map** usually shares the same geographic area of the overall map whereas an **insert** is normally an enlarged portion of the overall map. The submaps can be of different scales and projections.

2. Topology Handling

MACDIF supports the explicit coding of 2-cells by a sequence of segments forming a closed loop. An areal feature composed of these closed loops need not be contiguous

(i.e. it can be comprised of several physically separated loops) and can contain islands. Since each areal feature points to all its component segments and each segment is shared by two polygons, this scheme would require the duplication of the segment numbers in the file.

An alternate method of coding a 2-cell in MACDIF allows the sharing of a segment by polygons. With this method, the basic topological element is a topological segment (a 1-cell) which is a line associated with a topological definition and terminates at two **nodes** (0-cells). A **Topological Definition Section** in MACDIF contains the following information related to topological segments:

- An identification number for the beginning and end nodes. These are arbitrary numbers, not segment numbers, used only to represent this relationship. As such, a node is not a real feature in MACDIF.
- The left and right areas (2-cells). A topological segment entirely surrounded by an area is indicated by the lack of a pointer to the right area. Such a segment is not connected to other segments and is floating inside an area. Note that a segment having the same area on both sides is not necessarily floating.
- Other segments sharing the same beginning and end nodes.

A point-oriented feature *cannot* become a node, nor can a boundary segment for an areal feature become a 2-cell through the use of the Topological Definition Section. This method of representing topology is based entirely on the topological segment. A feature definition for the features on both sides of the segment can still be included in the MACDIF file, but they no longer point to segments and hence do not contain enclosure information which is associated only with segments. It is therefore possible to indicate that a segment is entirely enclosed by a polygon, but not possible to indicate that a polygon is an island surrounded by a larger polygon.

3. Coordinate System

The coordinates can be expressed in absolute, relative, differential, or incremental forms. An **absolute coordinate** measures the displacement from the origin of the coordinate system. A **relative coordinate** measures the displacement from the previous coordinate. A **differential coordinate** is a relative coordinate scaled. An **incremental coordinate** measures a displacement of fixed-size steps.

Any linear units of measurement can be used for the coordinates, but the conversion factor to metres must be supplied. Angular measurements are also supported to allow the use of latitudes and longitudes. Moreover, horizontal and vertical coordinates can be in different units. The number of significant digits for a coordinate can vary and is practically unlimited.

An interesting feature of MACDIF is that absolute and relative coordinates are normalized in that the rectangular map is regarded as a unit square. Each coordinate is hence a decimal number ranging from zero to one and transformation parameters carried

with the MACDIF file can be used to map these normalized coordinates to other systems such as the ground. The advantages of this scheme are:

- A different scale factor applied to the x and y coordinates will result in plots of various dimensions. This is the representation scheme used by most device-independent graphics software.

- The number of significant digits is optimized for the size of the area mapped.

4. Projections Defined

The parameters of the following 16 projections are defined in the MACDIF format: Mercator, General Transverse Mercator, Universal Transverse Mercator, Hotine Oblique Mercator, Oblique Mercator Spherical, Albers Equal Area Conic, Lambert Conformal Conic, Polyconic, Orthographic, Stereographic, Universal Polar Stereographic, Lambert Equal Area, Azimuthal Equidistant, Gnomonic, Azimuthal Equal Area and Plate Carrée (Simple Latitude/Longitude).

Other projections to be defined in later versions include: Space Oblique Mercator, Miller Cylindrical, Equidistant Cylindrical, Bonne, Lambert Equal Area Conic, Bipolar Oblique Conic Conformal, Modified Polyconic, Equidistant Conic, Miller Prolated Stereographic, Van Der Grinten, Sinusoidal and Eckert.

5. Spatial Objects

The MACDIF standard supports the following spatial objects:

- **Real Segment** - The positional definition of a point, a circle, a circular arc, a connected series of points to form a line, or a connected series of lines and arcs. Many line segments can be used to form a line features or form the perimeter of an areal feature. A segment can also be shared by different features.

- **Virtual Segment** - A segment used to define an imaginary line and used to close a partially bounded areal feature such as a lake with an opening at the place where it meets the mouth of a river. It is organized like a real segment.

- **Neat Line** - A special segment forming the border of the map.

- **Point-Oriented Feature** - A feature associated with a segment containing a single point. The definition of a point-oriented feature, similar to other features, contains the following items:

  - Feature Number - A unique numeric code for the feature.
  - Feature Type and Feature Sub Type - A feature code assigned according to the feature type table identified for the sub-map.

- Attribute Values - A record of attribute values associated with this feature. The values can be in any form and their interpretation is contained in the feature type table.
- Feature Class - A code indicating if this is a point-oriented-, a line-, an areal-, or a special feature. Objects like compass roses can be defined as special features.
- Feature Reference Position - A reference point for the feature. For a point-oriented feature, it is the location of the point. Otherwise, it is a point roughly at the centre of line- and areal features.
- Bounding Rectangle - A rectangle defined by the minimum x, minimum y, maximum x and maximum y of the coordinates.
- Annotative Name - A text string containing the name of the feature.
- Alternate Name - An alternate annotative name, for example in another language.
- Data Group - The numeric identification of the data group to which this feature belongs.
- Source Identification - A declaration of data source.
- Quality Identification - A declaration of data quality.
- Segment - The segment containing the coordinate of the point.
- Comment - A text string containing comments about this feature.

- **Line Feature** - A feature associated with one or more line segments. It is organized like a point-oriented feature.

- **Areal Feature** - A feature associated with one or more segments forming a closed area. It is organized like a point-oriented feature except that the segments can form one or more loops. Each loop is assigned a nesting level number.

- **Topological Segment** - A segment (1-cell) topologically related to nodes (0-cells) and areas (2-cells). Such a segments is associated with the beginning and end nodes, the left and right areas and all other segments sharing the beginning and end nodes.

6. Coding Method

A MACDIF file starts with an **Administrative Header** containing information such as the name of the file, the scale, the coverage area and statements about data source and data quality. This is followed by one or more **Map Sections** containing information about the various maps in the file.

Each Map Section starts with an optional **Sub Header** which is similar in structure to the Administrative Header. The use of the Sub Header to alter the parameters stored in the comprehensive Administrative Header to suit the map it describes. The first map in the file is the overall map while the other maps are the sub-maps.

After the Sub Header of each Map Section, the sections on transformation, features, segments, topology, symbology and associated information, are stored. The section on transformation contains the parameters required to map from the internal normalized

coordinate system to the world system. The section on features contains all the feature definitions and that on segments contains all the segment definitions.

The z-coordinates are often much shorter than the horizontal coordinates. In order to allow more compact coding of a MACDIF file, the z-coordinates can be stored as a separate string after the horizontal coordinates. The vertical measurements stored in this format are called the **associated z** and is particularly suitable for contours because one single z value can be associated with the entire line.

MACDIF, like other international standards for data communication, uses two levels of data description. The higher level provides a logical description of the format using a notation very often used to describe the syntax of a programming language. MACDIF adopted the **Abstract Syntax Notation (ASN)** defined by the International Standards Organization (ISO). For example, the following describes the partial format of a MACDIF file:

```
Digital-Map-or-Chart  ::=   SEQUENCE{ Map-Header, Map-Def-Section }
Map-Header            ::=   SEQUENCE{
                                  [1] Data-Set-ID, ... }
Map-Def-Section       ::=   SEQUENCE OF Map-Definition
Map-Definition        ::=   SEQUENCE {
                                  [1] Map-Sub-Header OPTIONAL, ...}
```

Definition one above shows that a MACDIF file is composed of a Map Header and a Map Definition Section. A more detailed explanation of the structure of the Map Header is given in definition two which indicates that the first item of a Map Header is the Data Set Identification Code. A later entry (not shown) in the ASN description will define in more detail how the Data Set Identification should be entered. Definition three shows that the Map Definition Section is actually a repeated sequence of Map Definitions each of which starts with an optional Map Sub Header as indicated in definition four.

The high level ASN notation gives a precise description of the data structure used in the interchange format. However, it is not the format used in the communication of data, which is the concern of the lower level data description. Several standards are available for the compact encoding of data for physical communication. The one used by MACDIF is based on the ISO Standard 8825 called the ASN.1 Encoding Rules. A string of bytes results from this encoding and some of them are printable ASCII characters and others are not.

The encoding standard does not dictate the medium for communication. In other words, the MACDIF format can be used for the interchange of digital data through tapes, disks, or a computer network. Naturally, even lower standards are required to handle data transfer through each of the media.

7. Strengths and Weaknesses

There are several features that make MACDIF different from the other standards examined in this report. They are summarized as follows:

- **Device Independence** - The strength of the multi-layered communication model proposed by the ISO and adopted by MACDIF is that it can be modified to suit different hardware and software environments.

- **Unlimited Precision** - The lengths of the numbers, coordinates and text strings used in MACDIF are unlimited.

- **Compact Encoding** - The ISO Standard 8825 for the encoding of the ASN syntax offers several methods of data compaction. MACDIF itself also provide meaningful defaults to many items so that an entry is only required in the file to override the prevailing value.

- **Modular data structure design** - The separation of segment-, feature- and topology definitions makes it easier to extend the format.

- **Customized for Maps and Charts** - MACDIF is designed to handle both mapping and charting data. It supports an extensive list of 16 projections (12 more to be added), 18 reference ellipsoids, 46 geodetic datums and 20 hydrographic tidal datums. Moreover, it can reference a number of national feature code tables. Other features particularly useful for mapping and charting are bilingual names, non-contiguous areal features and map inserts.

One of the weaknesses of MACDIF is in topology definition. A node cannot be a feature and nested polygons are not handled properly. Another area that needs development is the definition of the symbol tables. The complexity of MACDIF, which is one of its strengths, is perhaps one of its weaknesses due to the level of expertise required to understand and use the standard.

8. Reference

Canadian Hydrographic Service (1988). *Specification of the Map And Chart Data Interchange Format (MACDIF)*, Version 2, March, Canadian Hydrographic Service, Department of Fisheries and Oceans, Ottawa.

E. Standard Interchange Format (SIF)

1. Introduction

The SIF interchange format was designed by the Intergraph Corporation of Alabama for the exchange of 2-D and 3-D graphics data between systems via magnetic tapes. It is sometimes referred to as the Intergraph Standard Interchange Format (**ISIF**).

A SIF interchange file set consists of a SIF environment file and a SIF data file. The **SIF environment file** is an ASCII, sequential and variable-length file that controls the processing of the SIF file. It specifies, for example, the name of the output SIF file,

the format of integer numbers and the organization of the tape files. The **SIF data file** on tape contains the definition of the graphic data and is often referred to as the **SIF file.** It is a fixed-length, sequential file that can be in text (ASCII or EBCDIC) or binary format as indicated in the environment file.

Programmes are available from Intergraph to convert text data files to binary form and vice versa. This allows users to enter interchange commands using a text editor to generate the data files in text form and then choose either the text format or the more compact binary form to transmit the results.

A SIF file consists of either 2-D or 3-D **graphic elements**. An element can belong to one of 63 **levels**. A **range**, commonly known as the bounding rectangle and is defined by the minimum x and y and maximum x and y coordinates that can be associated with each graphic element. Several elements can also be formed into one of 65535 **graphic groups**. These groups can then be manipulated as if they were a single graphic element.

Attribute values can be attached to graphic elements through an ASSOCIATED VALUE entry before the graphic element definition. For this to be done properly, the Intergraph DMRS database utility should be used since the SIF file does not contain the data dictionary which defines how the fields are used in the attribute record. An ASCII description can also be attached to each graphic group through the use of a GRAPHIC ASSOCIATION DESCRIPTOR.

Point symbol (also called **cell**) definitions can be included in the SIF file. These symbols can be expanded from an existing Intergraph **cell file** during an export operation and can be inserted into a cell file during an import. The name of the cell file referenced is contained in the environment file. **Line Symbology (pattern)** can also be specified in the SIF file. However, pattern definitions reference the cells.

2. Topology Handling

SIF does not recognize topological entities except a polygon. Enclosure relationship are encoded by the use of holes.

3. Coordinate System

The coordinates in a SIF file are either signed or unsigned 32-bit integers, with units defined in the environment file as inches, feet or cm. Signed integers can range from -2147483648 to 2147483647 and unsigned integers can range from 0 to 4294967295. The origin of the coordinate system can be changed at any point in the SIF file.

4. Spatial Objects

The following geometric objects are supported by SIF:

- **Line string** - An open line defined by two or more points, or a closed line representing a solid polygon or a hole. Line strings must not exceed 101 points, but longer lines can be broken into segments of at most 101 points each. It is different from a curve in that a line string is not smooth.

- **Circle** - A circle defined by three points or "centre/radius" combination. A circle can be a hole.

- **Circular arc** - A part of a circle defined by three points or "centre/two points on edge/arc direction" combination.

- **Symbol Definition** - A graphic symbol (cell) defined by a symbol name (cell name) and its coordinates. These symbols are extracted from or to be placed into the Intergraph cell file.

- **Symbol** - A symbolized point defined by a position, a scale, a rotation and a symbol name (cell name). Symbol definitions are carried in the SIF data file or defined in a cell file.

- **Curve** - An open, smooth line passing through the given vertices. The number of vertices must be at least three but at most 101. A curve requiring more than 101 vertices must be broken into pieces of at most 101 vertices.

- **Ellipse** - An ellipse defined by the centre, a vertex on the primary axis and one on the secondary axis. An ellipse can be a hole.

- **Elliptical arc** - A part of an ellipse defined by the centre, a vertex on the primary axis, a vertex on the secondary axis, a start angle, and a sweep angle.

- **Conic** - A conic element defined by at least three vertices. One that is longer than 101 vertices must be broken into pieces of at most 101 vertices.

- **Complex string** - A string composed of the following lines: open line strings, arcs, elliptical arcs, curves and conics. A complex string can be closed to form solid polygons or holes.

- **Complex surface** - A 3-D surface composed of basic surfaces such as cylinders and conics.

- **Complex solid** - A 3-D solid composed of basic surfaces,

- **Point string** - A string of up to 48 points that are either connected or disjoint.

- **B-spline curve** - An open or closed curve defined as a rational or non-rational B-spline curve. A rational B-spline can be used to define conic sections such as ellipses. The major difference between a B-spline curve and a curve is that the B-spline curve does not interpolate the given points. A close B-spline curve can form a solid polygon or a hole.

- **B-spline surface** - A 3-D surface defined as a rational or non-rational B-spline surface. The surface can be open or closed. If closed, the B-spline surface can form a solid object or a hole in space.

- **Circular truncated cone** - A 3-D pointed or truncated cone with circular top and base. The top and base circles can be rotated with respect to the centre line. The cone can be a surface or a solid.

- **Cylinder** - A 3-D cylinder defined by the center of the base, the radius of the base, the center of the cylinder top. The top and base circles can be rotated with respect to the centre line. The cylinder can be a surface or a solid.

- **Text line** - A string of at most 255 characters. The font, origin, the height, the width and the angle of the string can be specified.

- **Paragraph** - A block of text defined by the origin, the height, the width and the angle of the strings in the paragraph.

- **Graphic Group** - A compound object which is composed of several graphic elements. A group is associated with a number ranging from 1 to 65535 and a GRAPHIC ASSOCIATION DESCRIPTOR containing a 1000-character comment string.

5. Coding Method

A SIF text data file is an unlabeled, 9-track, odd-parity tape of 800 to 1600 bpi. Logical records in the file are fixed-length from 20 to 80 bytes long each and the block size is fixed at 1024 bytes. The length of the logical records is defined in the environment file. A binary version of the SIF file has the same block size, but containing variable-length logical records. Multiple SIF files can reside on the same tape.

The SIF text data file consists of a series of commands up to 1024 characters long and can span over several lines (logical records). A command starts with a three-character keyword. This is followed by a slash "/" and other descriptor data each composed of a two-character keyword, an equal sign "=" and one or more data values. For example, the command to generate a circle with centre (10,20) and radius 40 is

CIR/CE=10,20,RA=40

If present, the DRAWING IDENTIFICATION COMMAND is the first command in the SIF file. It defines the drawing, the file name, the date the file was created, the graphic mode (2-D or 3-D), the range of the drawing file, the drawing units and a free-field identification. The order of most other commands is not important. There are numerous examples given in the Intergraph Implementation Guide to show the proper way of entering the interchange commands.

6. Strengths and Weaknesses

Of all the interchange formats examined in this report, SIF is the only one that supports a command language complete with keywords and qualifiers. The advantages of this format are that:

- It allows the user to create an interchange file interactively using a text editor. This is particularly attractive if an Intergraph user receives from another system a short file in text form containing simple graphic objects to be imported into the Intergraph system. This short text file does not have to conform to any interchange standard. As long as it can be processed by a text editor and contain well defined objects, it can be easily modified to take the form of a SIF text data file.

- It is more readable although it still needs an experienced user to interpret its contents.

SIF is rich in geometric objects but lacks topological objects. The only topological entity supported is a closed polygon with optional holes. There is also a lack of topographic objects such as contours and spot heights. One way is to transfer spot heights as a text in which the spot height label is entered as a text string and the height is encoded in the 16-bit integer, unique identification. As for contours, they can be represented by a 2-D line with the contour height carried in the numeric identification. There is no explicit fields for feature codes, but one or more numeric identifiers can be assigned to an element through the use of the IDENTIFIER command.

The transfer of attribute information using SIF is not convenient unless both parties use the Intergraph DMRS database files. Another weakness is that a text string cannot be bound to a curve.

7. Reference

Intergraph (1986). *Standard Interchange Format (SIF) Command Language Implementation Guide (8.8.2)*, July 1, Huntsville, Alabama.

APPENDIX III

EXAMPLES OF DATABASE DOCUMENTATION

A. EXAMPLE OF DATA QUALITY DOCUMENTATION: QUALITY REPORT FOR DANE COUNTY SOIL SURVEY, DANE COUNTY, WISCONSIN, U.S.A.

B. EXAMPLE OF SUMMARY DATABASE DOCUMENTATION:
LAND INFORMATION NEW ZEALAND.